Performance Contracting in After-Sales Service Supply Chains*

Sang-Hyun Kim, Morris A. Cohen, and Serguei Netessine†
The Wharton School, University of Pennsylvania
Philadelphia, PA 19104

January 2006, Revised August 2006.

Abstract

Performance-based contracting is reshaping service support supply chains in capital intensive industries such as aerospace and defense. Known as "power by the hour" in the private sector and as performance-based logistics (PBL) in defense contracting, it aims to replace traditionally used fixed-price and cost-plus contracts in order to improve product availability and reduce the cost of ownership by tying a supplier's compensation to the output value of the product generated by the customer (buyer).

To analyze implications of performance-based relationships, we introduce a multitask principal-agent model to support resource allocation and use it to analyze commonly observed contracts. In our model the customer (principal) faces a product availability requirement for the "uptime" of the end product. The customer then offers contracts contingent on availability to n suppliers (agents) of the key subsystems used in the product, who in turn exert cost reduction efforts and set spare parts inventory investment levels. We show that the first-best solution can be achieved if channel members are risk-neutral. When channel members are risk-averse, we find that the second-best contract combines a fixed payment, a cost-sharing incentive and a performance incentive. Furthermore, we show how these contracts evolve over the product deployment life cycle as uncertainty in support cost changes. We show, in particular, that when the customer is less (more) risk-averse than the suppliers, the performance incentive increases (decreases) while the cost sharing incentive decreases (increases) with time. Finally, we illustrate the application of our model to a problem based on aircraft maintenance data and show how the allocation of performance requirements and contractual terms change under various environmental assumptions.

*The authors are grateful to the seminar participants at the Wharton School, Cornell University, the University of Texas at Dallas and the University of Washington in Seattle for helpful discussions. The authors would also like to acknowledge comments received from the departmental editor Ananth Iyer, the associate editor and the two anonymous referees as well as input from L. Gill, S. Gutierrez, M. Lebeau and M. Mendoza, who provided valuable information concerning current practices. Finally, the authors are grateful for the assistance of Ashish Achlerkar of MCA Solutions, who provided valuable assistance in testing the model and providing access to a real-world data set. This research was partially funded by a grant from the National Science Foundation (Grant 540075-8001).

†Email: shkim@wharton.upenn.edu, cohen@wharton.upenn.edu, netessine@wharton.upenn.edu

1 Introduction

Support and maintenance services continue to constitute a significant part of the U.S. economy, often generating twice as much profit as do sales of original products. For example, a 2003 study by Accenture (see [1]) found that $9B in after-sales revenues produced $2B in profits for General Motors, which is a much higher rate of profit than its $150B in car sales generated over the same time period. According to the same study, after-sales services and parts contribute only 25% of revenues across all manufacturing companies but are responsible for 40-50% of profits.

Since after-sales support services are often provided and consumed by two different organizations (i.e., the OEM and the customer), the issue of contracting between them becomes important. While contracts for maintenance services of simpler products (electronics, automobiles) involve fixed payments for warranties, there are many instances of complex systems that require more sophisticated relationships between service buyers and suppliers. For example, in capital-intensive industries such as aerospace and defense, significant uncertainties in cost and repair processes make it very hard to guarantee a predetermined service level or quote a price for providing it. Therefore maintenance support in these industries typically involves cost-sharing arrangements, which include fixed-price and cost-plus contracts. Under the former, the buyer of support services pays a fixed fee to the supplier to purchase necessary parts and support services; under the latter, the buyer compensates the supplier's full cost and adds a premium.

Through our work with aerospace and defense contractors we have observed a major shift in support and maintenance logistics for complex systems over the past few years. Performance-based contracting, a novel approach in this area, is replacing traditional service procurement practices. This approach is often referred to as Power by the Hour or Performance-Based Logistics (PBL) in, respectively, the commercial airline and defense industries. The premise behind performance-based contracting is summarized in the official Department of Defense (DoD) guidelines (Section 5.3 in [11]): "The essence of Performance Based Logistics is buying performance outcomes, not the individual parts and repair actions... Instead of buying set levels of spares, repairs, tools, and data, the new focus is on buying a predetermined level of availability to meet the [buyer's] objectives." In 2003, the DoD issued Directive 5000.1 [10] which "*requires* program managers to develop and implement PBL strategies that optimize total system availability." Hence, all future DoD maintenance contracts are mandated to be performance-based.

A critical element of performance-based contracting is the clear separation between the buyer's expectations of service (the performance goal) and the supplier's implementation (how it is achieved). In the words of Macfarlan and Mansir [20], "The contract explicitly identifies what is required, but the contractor determines how to fulfill the requirement." As a consequence of this flexibility, PBL contracting should promote new and improved ways to manage spare parts inventory, reduce administrative overhead, negotiate contracts, and make resource allocation decisions. For example, under the traditional cost-plus contract, the supplier of a service must truthfully report its detailed cost structure to the buyer in order to determine which expenses are eligible for reimbursement. Under a PBL arrangement, the supplier does not have to share cost information at this level of detail. Moreover, the product buyer no longer directly manages or possibly even owns resources such as the inventory of spares. Finally, in the long run suppliers may find it in their interest to invest in designing more reliable products and more efficient repair and logistics capabilities.

Not surprisingly, such a radical change in the approach to contracting has caused confusion among suppliers of after-sales support services. The academic literature, however, offers little guidance with respect to how such contracts should be executed. In this paper we aim to take a first step towards filling this void by proposing a model of performance-based contractual relationships that arise in practice when procuring repair and maintenance services. We embed a classical single-location spare parts inventory management problem into a principal-agent model with one principal (representing the end customer), and multiple interdependent agents (representing suppliers of the key product subsystems). Each agent performs two tasks that are subject to moral hazard: spare parts inventory management and cost reduction. We use this model to analyze three types of contracts (and any combination thereof) that are commonly encountered in aerospace and defense procurement and high technology industries, namely, fixed-price, cost-plus and PBL. In analyzing these contracts we analyze the following questions: (1) what is the optimal combination of contractual levers that achieves the best possible outcome for the buyer? (2) how should a performance requirement for the final product be translated into the performance requirements for the suppliers who provide critical components? and (3) how should the risk associated with the maintenance of complex equipment be shared among all supply chain members?

We show that, if suppliers' decisions are observable and contractible, the contract that achieves the first-best solution is a nonperformance arrangement that combines partial cost reimbursement with a fixed payment. If supplier actions are unobservable and the parties are risk-neutral, we show that the first-best solution can still be achieved using a contract that combines a performance incentive with a fixed payment (but no cost sharing). However, when even one of the parties is risk-averse, the first-best solution cannot be achieved. In this case, we show that "pure" fixed-price, cost-plus, or performance-based contracts (or any pair-wise combination of them) are not suitable because they do not provide the necessary incentives. Thus, we show that the second-best contract involves all three elements: a combination of a fixed payment, a cost sharing payment and a performance-based payment.

For any such contract proposed by the customer, we find analytically optimal decisions for all suppliers. Unfortunately, the buyer's problem neither is well-behaved nor admits tractable analytical solutions (the latter is true even in the centralized supply chain). Using a combination of analytical results for special cases and numerical analysis performed on a data set that is representative of a supply chain supporting a fleet of military airplanes, we obtain insights into the structure of the optimal contract. In particular, we study the sensitivity of the optimal contract to cost uncertainty and infer that, when the principal is less (more) risk-averse than the suppliers, the performance incentive increases (decreases) while the cost sharing incentive decreases (increases) as time progresses. Finally, we analyze the impact of problem parameters on contractual terms, performance, and profitability.

To the best of our knowledge, this paper represents the first attempt to embed the after-sales service supply chain model into the principal-agent framework in which the supply chain members behave in a self-interested manner. Our results are consistent with the observed practice of using multiple contract types whose mix evolves over time. Finally, the model framework introduced here can be implemented in conjunction with more detailed supply chain models to support contract negotiations and long-term strategic analysis. The rest of the paper is organized as follows. After a brief review of related literature in Section 2, we present modeling assumptions and notations in Section 3. In Section 4 we analyze the first-best solution as well as derive solutions for the general second-best case. In Section 5 we analyze special cases, beginning with the risk-neutrality assumption, then we study an environment in which the suppliers' actions are partially observable, and finally we study a situation with one supplier. Two numerical examples are presented in Section 6, including the one based on the aforementioned military aircraft data set. Finally, in Section 7 we discuss managerial implications of our study.

2 Literature Review

Two distinct models are blended together in our paper: a classical inventory planning model for repairable items, well known in operations management, and the moral hazard model that has been an area of active research in economics. The theory of repairable parts inventory management dates back to the 1960s when Feeney and Sherbrooke [13] introduced a stochastic model of the repairable inventory problem whose steady-state solution relies on the application of Palm's Theorem. Sherbrooke's METRIC model (Sherbrooke [28]) and its extensions, such as Muckstadt [23] and Sherbrooke [29], established the basic modeling framework and heuristic optimization algorithms for allocating inventory resources in multi-echelon, multi-indentured environments. Subsequent models have led to notable success in enabling the management of multimillion-dollar service parts inventory resources in both commercial and government applications (e.g., see Cohen et al. [7] for a discussion of a successful application of multi-echelon optimization by IBM's service support division). Research in this area has largely focused on improving computational efficiency and on incorporating more realistic modeling assumptions, such as allowing for capacitated supply or nonstationary demand processes. For a recent comprehensive account of developments in this field, see Muckstadt [24], who reviews the underlying theory, Sherbrooke [30], which focuses on aerospace and defense industry applications, and Cohen et al. [6], which introduces a modeling framework that has been used to guide the development of state-of-the art software solutions in various industries.

In brief, repairables inventory models are concerned with finding optimal (cost-minimizing) inventory stocking targets for each product component subject to an overall service constraint. Service (performance) requirements can be defined in terms of either item fill rates or end product availability (i.e., system "uptime"). The latter is the preferred choice in aerospace and defense environments, and we adopt it in our paper (for a discussion of comparison of these metrics, see Sherbrooke [28]).

Numerous papers study the principal-agent models, and comprehensive review can be found in Bolton and Dewatripont [2]. The building block for our paper is the moral hazard model in which actions of agents (suppliers) are unobservable to the principal (customer). Moreover, our model includes elements of multitasking (Holmström and Milgrom [14]), because the two decision variables for suppliers, the cost reduction effort and the inventory position, interact with each other. An additional complication is the presence of multiple agents whose decisions together impact the performance constraint that the principal faces. There are a number of economics papers discussing cost reimbursement contracts in the presence of moral hazard. For example, Scherer [27] considers optimal cost-sharing and the impact of risk aversion in defense procurement. McAfee and McMillan [22] presents a model in which firms bid for government contracts under significant cost-related risks. Inspired by this research, we allow for risk aversion and study cost-plus and fixed-price contracts in the context of after-sales support and compare them with performance contracts. In the operations management literature, work of So and Tang [32] is closely related to our paper in that they also consider outcome-based reimbursement policies, but their focus is on the health-care industry.

Incentive alignment in supply chains through contracts has been a topic of great interest in operations management over the past decade (see Cachon [3] for a comprehensive survey). Recently, the role of information asymmetry has received considerable attention both in the adverse selection setting (representative articles include Corbett [8], Iyer et al. [15], Lutze and Ozer [19] and Su and Zenios [33]) and in the moral hazard setting (for example, see Plambeck and Zenios [26], Chen [4] and Plambeck and Taylor [25]). The current paper contributes to this growing area as well.

As is evident from our survey, although there is voluminous literature on repairables inventory management, to date this stream of research has been confined to single-firm models and hence does not address issues that arise in decentralized supply chains. Furthermore, although the extensive literature in economics aims to model contractual relationships among different parties, it does not address the complexities of repair and maintenance contracting environments. To our knowledge, our paper is the first to put a repairables inventory model into the decentralized framework and to study the issue of contracting in after-sales service supply chains.

3 Modeling Assumptions

The principal is the customer for N identical assembled products ("systems," which can be airplanes, computers, manufacturing equipment, etc.) Each system is composed of n distinct major parts ("subsystems" which, in the case of an airplane, can represent avionics, engines, landing gear, weapons systems, etc.), each produced and maintained by a unique supplier. We use subscript 0 to denote the customer and subscript i for subsystem supplier i, $i = 1, 2, ..., n$. We ignore the indenture structure in the subsystem's bill of materials, treating each subsystem as a single composite item. In the following subsections, we describe the repair process and supplier cost structure, explain how risk aversion is represented, define the performance measure, specify contract terms, and derive the utilities of the customer and the suppliers.

3.1 Repair Process

Failure of the subsystem i is assumed to occur at a Poisson rate λ_i, independently from failures of other subsystems. Each supplier maintains an inventory of spares and a repair facility. A one-for-one base stock policy is employed for spares inventory control. That is, a failed unit is immediately replaced by a working unit (if it is available) from the supplier's inventory. If a replacement is unavailable, a backorder occurs, and the affected system becomes inoperable. As a result, downtime in any subsystem leads to downtime of the system. Upon failure, the defective unit immediately goes into the repair facility, modeled as an M/G/∞ queue. We assume ample capacity (i.e., infinite number of servers) which is an idealization of reality, but it is considered a reasonable approximation (Sherbrooke [30]). This assumption leads to the desirable property that repair lead times of different items are independent. It takes, on average, L_i time units to repair the subsystem, and once the task is completed the subsystem is placed in the supplier's inventory. Forward and return transportation lead times are incorporated into the repair lead time and are assumed to be independent of the customer location (Wang et al. [35] relax this assumption).

The number of backorders of subsystem i, B_i, is a random variable that is observed at a random point in time after steady state is reached. Supplier i chooses a target spare stocking level s_i for

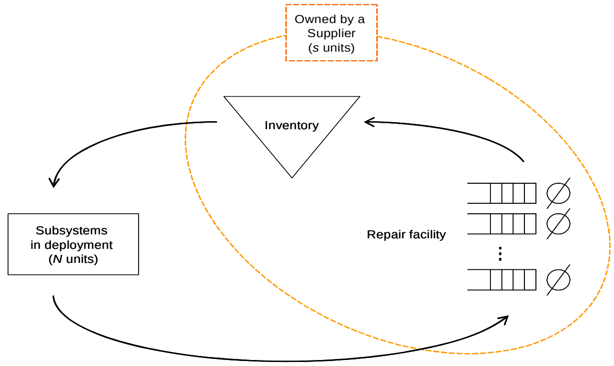

Figure 1: Closed loop cycle for repairable items.

subsystem i. B_i and s_i are related to each other through $B_i = (O_i - s_i)^+$, where O_i is a stationary random variable representing the repair pipeline (on-order) inventory. Palm's Theorem states that O_i is Poisson-distributed for any repair lead time distribution, with the mean $\mu_i \equiv \lambda_i L_i$ (Feeney and Sherbrooke [13]).

It is important to point out that the repair process forms a closed-loop cycle. As the subsystems are typically very expensive and their lifetimes are very long, we assume that no subsystem is discarded during the entire support period. Figure 1 illustrates this process. Thus there are a total of $N + s_i$ units of subsystem i in the supply chain, but only s_i of them are owned by the supplier. The fixed failure rate assumption is in fact an approximation, as the closed-loop cycle with finite population means λ_i is a function of the number of working units. However, this approximation is reasonable in our problem context because the condition $E[B_i \mid s_i] \leq \lambda_i L_i \ll N$ is satisfied in practice for most spare subsystems. This condition ensures that on average the number of subsystems being repaired at any given time is relatively small, and that the correction due to state-dependency can be safely ignored. Indeed the fixed failure rate assumption is used in virtually all analytical models in the literature, beginning with the seminal paper by Sherbrooke [28].

Although the Poisson distribution arising from Palm's Theorem is appealing, it turns out that working with integer-valued random variables O_i and B_i as well as the discrete decision variable s_i complicates our analysis significantly. This happens because of complexities resulting from the game-theoretic situations associated with various contracting options. In particular, deriving tractable mathematical expressions to gain insights into firms' behavior becomes prohibitively complex. For this reason, we depart from the usual discrete distribution assumption and model O_i, B_i, and s_i as continuous variables. This approach is reasonable in our context since each unit of a supplier's inventory represents a composite of the various components associated with their particular subsystem. Such aggregation results in sufficiently high values for μ_i so that normal approximation for the Poisson distribution can be applied (see Table 2 for a range of μ_i used in a numerical example presented in Section 6.2). However, the normality assumption is not essential – all our results in the paper are derived for an arbitrary distribution. To this end, we let O_i be distributed continuously with cdf F_i and pdf f_i, which have nonnegative support $[0, \infty)$ with $F_i(0) \geq 0$. The distribution of B_i is obtained from $\Pr(B_i \leq x \mid s_i) = \Pr(O_i \leq x + s_i)$. Furthermore, $E[B_i \mid s_i] = \int_{s_i}^{\infty} [1 - F_i(x)]\, dx$ so we obtain

$$dE[B_i \mid s_i] / ds_i = -1 + F_i(s_i) \leq 0, \qquad d^2 E[B_i \mid s_i] / ds_i^2 = f_i(s_i) \geq 0. \tag{1}$$

Hence, we see that the expected backorder is decreasing and convex in s_i.

3.2 Supplier Cost

Supplier i's total cost to maintain its subsystem, C_i, has fixed and variable components. The fixed cost contains an additive stochastic term ε_i. The expected fixed cost can be normalized to zero without affecting any of our results because it does not play a role in determining optimal supplier decisions and contract terms, as will become evident in the next section. The variable cost is equal to the unit cost of a spare subsystem, c_i, times s_i, the number of units in the base stock. ε_i represents the uncertainty in total cost that is beyond the supplier's control, and is assumed to have zero mean and a finite variance. $\{\varepsilon_i\}$ are assumed to be uncorrelated across suppliers, i.e., $\text{Cov}[\varepsilon_i, \varepsilon_j] = 0$ for $i \neq j$. Furthermore, we assume that $\text{Cov}[\varepsilon_i, B_i] = \text{Cov}[\varepsilon_i, B_j] = 0$ holds for all $i \neq j$. The uncertainty in the unit cost is assumed to be negligible compared to ε_i. This assumption is based on our discussions with practitioners who indicated that the uncertainty with respect to fixed costs is of greater importance during the support stage. The unit cost uncertainty may be significant during the product development stage, but we do not model it in this paper. In addition, we assume that c_i and the distributions of ε_i and B_i are common knowledge.[1]

The total cost of support can be reduced by the dollar amount a_i, which is interpreted as the supplier's discretionary effort. Hence, $C_i = c_i s_i - a_i + \varepsilon_i$. By exerting effort the supplier incurs disutility $\psi_i(a_i)$, which is convex increasing ($\psi_i'(a_i) > 0$, $\psi_i''(a_i) > 0$) and $\psi_i(0) = 0$. In the sequel, we assume a quadratic functional form $\psi_i(a_i) = k_i a_i^2/2$ with $k_i > 0$. This assumption does not fundamentally change the insights of our model, while generating compact expressions, and for this reason it is commonly used in the literature (see, for example, Chen [4]). As is customary in the literature, we take the accounting convention that C_i is observable by the customer and is the basis of reimbursement (see

[1]That the unit cost is known to the customer is plausible in the defense industry, as most of the current PBL contracts apply to existing subsystems whose unit cost had to be revealed under pre-PBL relationships. In traditional defense contracting, the DoD negotiates the price of a spare part or a subsystem based on the reported unit cost.

Laffont and Tirole [18] p. 55).[2]

The crucial distinction between the supplier's actions a_i and s_i is the way each variable contributes to the performance outcomes; the backorder is influenced by s_i only, since $B_i = (O_i - s_i)^+$, while the total cost is affected by both decision variables a_i and s_i such that $C_i = c_i s_i - a_i + \varepsilon_i$. This interaction creates asymmetry in how the suppliers' actions influence outcomes B_i and C_i. Raising a_i reduces the total cost but has no impact on availability, which is driven by component reliability and repair processes. That said, raising s_i improves availability but incurs a higher cost. The latter is the classical cost-availability trade-off seen in repairables inventory models. We do not consider an alternative formulation, whereby supplier effort impacts product reliability and/or repair capabilities (thus impacting λ_i and L_i).[3]

3.3 Risk Aversion

We assume that all members of the supply chain are risk-averse with expected mean-variance utility

$$E[U_i(X)] = E[X] - r_i \mathrm{Var}[X]/\,2. \tag{2}$$

The constant $r_i \geq 0$ is the risk aversion factor, representing the member i's inherent attitude towards uncertainty. Great uncertainties that pervade product development, production, and maintenance mean significant risks for the firms, and their risk-averse perspective is commonly observed (see Scherer [27] for discussion and references). The larger the value of r_i, the more risk-averse a firm is, whereas risk neutrality is a special case with $r_i = 0$. This form of utility function has been widely used in recent operations management literature because of its tractability (Chen and Federgruen [5], Van Mieghem [34]).

3.4 Performance Measure

The performance metric in our problem is availability A_i, which is defined as the fraction of deployed systems that have a functional subsystem i at a random point in time. By this definition $A_i = 1 - B_i/N$ because a backorder of each subsystem results in a backorder for the entire system. Similarly, let us

[2]Note that the supplier is not compensated for his disutility of effort $\psi_i(a_i)$. With this convention, we model effectively assume that that the cost reduction effort a_i is the supplier's own discretionary decision and hence the customer does not subsidize the supplier's internal cost for it. In other words, the customer reimburses only the undisputable *direct* cost of maintenance that would withstand the scrutiny of a possible audit.

[3]Having fixed λ_i and L_i is a reasonable representation of reality in the defense industry. At present, most PBL contracts are awarded for existing systems whose subsystem specifications (hence reliability) and repair facilities (e.g., specialized equipment, tools) are already set and cannot be easily altered. Capacity/reliability decisions are typically made at the product development and initial deployment stages. Once they are made, improvements occur infrequently as the product re-engineering processes require long lead times and significant up-front investments.

define the system availability A_0 as the fraction of deployed systems that are fully functional. Unlike in the case of subsystems, this definition in general does not imply $A_0 = 1 - B_0/N$ where $B_0 = \sum_{i=1}^{n} B_i$, since it is possible that a system is down due to backorders of more than one subsystem. In fact, one can see that $1 - \sum_{i=1}^{n} B_i/N \leq A_0 \leq 1 - \max_i\{B_i\}/N$. However, a common assumption in the literature (for example, see Muckstadt [24]) is that the probability of two or more subsystems being down within the same system at any point in time is negligible. This is a reasonable assumption since the failures of deployed subsystems typically occur very infrequently. Thus the relation $A_0 = 1 - B_0/N$ holds under this assumption. As a consequence, we assume that all system failures are caused by single subsystem failure and there is no ambiguity in assigning accountability for system downtime to a specific supplier.

Because of the one-to-one correspondence between A_0 and B_0 resulting from the assumption, the system availability requirement $E[A_0 \,|\, s_1, s_2, ...s_n] \geq \widehat{A}_0$ (e.g., "expected system availability has to exceed 95%") is equivalent in our model to a system backorder constraint $E\left[B_0 \,|\, s_1, s_2, ...s_n\right] = \sum_{i=1}^{n} E\left[B_i \,|\, s_i\right] \leq \widehat{B}_0$. We call $\widehat{B}_0$ the system backorder target. Additionally, we assume that $\sum_{i=0}^{n} \mu_i > \widehat{B}_0$ in order to rule out the trivial case in which $s_1 = s_2 = ... = s_n = 0$ is optimal.

We note that our focus on performance incentives raises the question of how our definition of availability can be used to quantify performance. One approach would be to compute the backorder metric as the time average of the stationary backorder process (indexed by time t), and not the stationary variable itself. That is, the random variable $\widetilde{B}_i(\tau) \equiv \frac{1}{\tau}\int_0^{\tau} B_i(t)dt$ could be used, instead of B_i, as the performance metric. τ is the horizon over which the number of backorders are counted and averaged. The distinction between these two measures is inconsequential in a risk-neutral setting since the utilities of the customer and the suppliers are functions of expectation only, and $E[\widetilde{B}_i(\tau)] = E[B_i]$. For the case of risk aversion, however, these two measures diverge because $\mathrm{Var}[B_i \,|\, s_i]$ is independent of τ while $\mathrm{Var}[\widetilde{B}_i(\tau) \,|\, s_i]$ decreases with τ. The latter is a consequence of the ergodic property of $B_i(t)$. Therefore, if we were to adopt $\widetilde{B}_i(\tau)$ as our performance measure, uncertainty with respect to availability may become insignificant with sufficiently large τ. However, this is not a good representation of reality because the customer and the suppliers alike express major concerns about performance variability at any point of time rather than time-averaged performance. For example, the U.S. Air Force currently uses a confidence interval of availability as its performance metric, for managing the risk associated with mission readiness. The computation of this interval is based on the steady state variable B_i (Slay et al. [31]). Thus, in order to reflect current practice for computing performance, we choose availability defined in terms of steady state backorders as the appropriate performance measure.[4]

[4]For completeness, we have investigated the impact of choosing the alternative measure $\widetilde{B}_i(\tau)$ and found that none of the qualitative results in this paper change. See the Kim et al. [16].

3.5 Contract Terms and Utilities

The customer's payment (transfer) to the supplier i is comprised of three terms: (1) a fixed payment, (2) reimbursement for the supplier's cost, and (3) a backorder-contingent incentive payment. Specifically, it has the form

$$T_i(C_i, B_i) = w_i + \alpha_i C_i - v_i B_i, \tag{3}$$

where w_i, α_i, and v_i are the contract parameters determined by the customer. w_i is the fixed payment, α_i is the customer's share of the supplier's costs, and v_i is the penalty rate for each backorder incurred by the supplier. With $v_i = 0$ and $\alpha_i = 0$, we obtain a fixed-price (FP) contract; with $\alpha_i = 1$ and $v_i = 0$ we obtain a cost-plus (C+) contract with full reimbursement.

Under the assumptions we have laid out so far, supplier i who is given a contract $T_i(C_i, B_i)$ has the following expected utility:

$$\begin{aligned} E\left[U_i\left(T_i(C_i, B_i) - C_i\right) - \psi_i(a_i) \,|\, a_i, s_i\right] \quad = \quad & w_i - (1-\alpha_i)(c_i s_i - a_i) - v_i E\left[B_i \,|\, s_i\right] - k_i a_i^2/2 \\ & - r_i(1-\alpha_i)^2 \mathrm{Var}[\varepsilon_i]/2 - r_i v_i^2 \mathrm{Var}[B_i \,|\, s_i]/2. \end{aligned} \tag{4}$$

The first three terms together represent the expected net income of the supplier, while the fourth term is internal disutility for exerting cost reduction effort. The last two terms, respectively, represent risk premiums associated with cost and performance uncertainties. Similarly, the customer's expected utility is

$$\begin{aligned} E\left[U_0\left(-\textstyle\sum_{i=1}^n T_i(C_i, B_i)\right) \,|\, \{a_i, s_i\}\right] \quad = \quad & -\textstyle\sum_{i=1}^n \big(w_i + \alpha_i(c_i s_i - a_i) - v_i E\left[B_i \,|\, s_i\right] \\ & + r_0 \alpha_i^2 \mathrm{Var}[\varepsilon_i]/2 + r_0 v_i^2 \mathrm{Var}[B_i \,|\, s_i]/2\big). \end{aligned} \tag{5}$$

That is, the customer's utility is a function of her total expenditure only. Lastly, each supplier is assumed to have reservation utility determined from its external opportunities. Without loss of generality, we can be normalize its value to zero.

3.6 Sequence of Events

Our representation of the after-sales support relationship is based on the standard single-location, steady-state repairables model with a take-it-or-leave-it contract. We do not consider issues arising from repeated interactions between the customer and the suppliers in this paper.[5]

Under these assumptions of the model, the sequence of events is as follows. (1) the customer offers the suppliers take-it-or-leave-it contracts; (2) the suppliers accept or reject the contracts; (3) the suppliers who have accepted the contracts take cost reduction measures and set the base stock levels of their spares inventory; (4) realized costs and backorders are evaluated at the end of the contract horizon; and finally (5) suppliers are compensated according to the contract terms.

4 Analysis

In the performance-based contracting environment, neither the details of supplier cost nor how the supplier meets their performance objectives is revealed to the customer. Instead, each supplier is compensated based on his total realized cost C_i and his realized backorder level B_i. The fact that both of these contractible variables include random noise raises the issue of incentives. Since C_i and B_i are functions of the supplier's cost reduction effort a_i and base stock level decision s_i, the supplier can partially control the performance related to his subsystem and his compensation by setting a_i and s_i. However, he may choose (a_i, s_i) that are not optimal from the customer's point of view. For example, an opportunistic supplier may choose to minimize his own disutility of efforts by "shirking" (i.e., choosing low a_i and s_i), hoping that a fortuitous state of the world is realized. The customer's task is then to provide appropriate incentives through contract terms that would induce the supplier to make the desired decisions. The customer's objective is to maximize her expected utility (or minimize her negative utility) subject to the system availability requirement, or equivalently the backorder requirement. In the following discussions, we use the term "observable" to mean that a variable is both observable and verifiable and hence can be specified in a contract. We first present the benchmark case with complete observability, and then consider the private action case. This section concludes with a comparison of common contracting options.

4.1 First-Best Solution: Complete Observability of Suppliers' Actions

In this subsection we analyze the problem under the assumption that suppliers' actions $\{a_i, s_i\}$ are both observable, a situation often referred to as the first-best solution because the customer avoids incentive problems by dictating $\{a_i, s_i\}$ to the suppliers. This is the benchmark case against which we

[5]Due to uncertainties in fleet deployment schedules and future support budgets, the DoD is unwilling to sign long-term contracts (i.e., for the life of the program), and instead typically contracts on a shorter-term basis with annual adjustments. Suppliers typically conduct multi-period budget planning using a short-term, steady-state model on a rolling horizon basis. Therefore, making a single-interaction assumption is appropriate. Although pre-contractual bargaining or renegotiation may exist in practical situations, we do not formally model them in this paper.

can evaluate the efficiency of other contracts. The customer's problem is

$$
\begin{array}{lll}
(\mathcal{A}_{FB}) & \min\limits_{\{w_i,\alpha_i,v_i,a_i,s_i\}} & E\left[U_0\left(\sum_{i=1}^n T_i(C_i,B_i)\right) \mid \{a_i,s_i\}\right], \\
& \text{s.t.} & \sum_{i=1}^n E\left[B_i \mid s_i\right] \le \widehat{B}_0, \qquad \text{(AR)} \\
& & E\left[U_i\left(T_i(C_i,B_i)-C_i\right)-\psi_i(a_i) \mid a_i,s_i\right] \ge 0 \quad \forall i, \qquad (\text{IR}_i) \\
& & 0 \le \alpha_i \le 1, \quad a_i,s_i \ge 0 \quad \forall i.
\end{array}
$$

The expected utility expressions are given by (4) and (5). (AR) is the system availability requirement constraint expressed in terms of backorders, and (IR_i) is the individual rationality constraint that ensures supplier i's participation. As is typical in moral hazard problems, each (IR_i) constraint binds at the optimal solution. That is, the customer is able to extract all of the surplus from the suppliers by setting appropriate fixed payments $\{w_i\}$. Let θ be the Lagrangian multiplier of the customer's problem ($\mathcal{A}_{FB}$). The following proposition specifies the first-best solution.

Proposition 1 *When the suppliers' decisions are observable and contractible, the optimal contract specifies the following supplier decisions* (a_i, s_i)*:*

$$
\begin{aligned}
a_i &= 1/k_i, && (6)\\
s_i(\theta) &= F_i^{-1}\left(\max\{1-c_i/\theta,0\}\right), && (7)\\
\textstyle\sum_{i=1}^n E\left[B_i \mid s_i(\theta)\right] &= \widehat{B}_0. && (8)
\end{aligned}
$$

The solution $\{a_i^{FB}\}$*,* θ^{FB} *and* $\{s_i^{FB}\} = \left\{s_i(\theta^{FB})\right\}$ *is unique and is obtained by offering a non-performance-based, risk-sharing contract such that* $v_i^{FB} = 0$ *and*

$$\alpha_i^{FB} = r_i/\left(r_0 + r_i\right) \tag{9}$$

provided that r_0*,* $r_i > 0$*. Supplier* i*'s expected utility is zero, whereas the customer's expected utility is* $\sum_{i=0}^n \left(-c_i s_i^{FB} + \frac{1}{2k_i} - \frac{1}{2}\frac{r_0 r_i Var[\varepsilon_i]}{r_0+r_i}\right)$*.*

We note that $\{s_i^{FB}\}$ and θ^{FB} are determined simultaneously from equations (7) and (8). Let us first consider the optimal risk sharing of cost, represented by (9). It is a modified version of the Borch rule (see Bolton and Dewatripont [2]). To gain insights, it is useful to consider extreme cases. If $r_0 > 0$ and $r_i = 0$, i.e., if supplier i is risk-neutral but the customer is not, $\alpha_i = 0$. This outcome corresponds to a FP contract; since the customer is risk-averse whereas the supplier is not, the customer transfers all risks to the supplier. At the opposite end, consider $r_0 = 0$ but $r_i > 0$, i.e., only the customer is risk-neutral. In this case $\alpha_i = 1$, meaning that the C+ contract is used. Although it may sound counterintuitive that the C+ contract achieves the first-best solution, we should recall that incentives are not an issue in the current setting because the suppliers' actions are observable and contractible. The role of the C+ contract is merely to mitigate the suppliers' reluctance to participate in the trade (the (IR_i) constraint). The risk-neutral customer can absorb all risks without efficiency loss. When both r_0 and r_i are positive, the customer and the supplier i share the risk according to (9), i.e., based on the supplier's risk aversion relative to that of the customer. For the remaining case $r_0 = r_i = 0$ (which is not covered by Proposition 1), risk sharing is not an issue and there is an infinite number of (w_i, α_i, v_i) combinations which are optimal, i.e., the solution is degenerate.

We now focus on the customer's first-best expected utility in which there are three terms for each supplier. The first term $(-c_i s_i^{FB})$ is the cost of owning s_i^{FB} units in the supplier's spares inventory. The second term $1/2k_i$ is the net savings due to the supplier's cost reduction efforts. The last term $\frac{1}{2}\frac{r_0 r_i}{r_0+r_i}\text{Var}[\varepsilon_i]$ can be interpreted as the joint risk premium between supplier i and the customer and it is positive only if they are both risk-averse. It represents the inefficiency created by a trade-off between the customer's desire to protect herself from risk and to facilitate the suppliers' participation, which require some degree of risk-sharing through cost reimbursement.

Unlike cost risk $\text{Var}[\varepsilon_i]$, performance risk $\text{Var}[B_i \mid s_i]$ poses no trade-off between the customer and the suppliers; it can be eliminated by setting $v_i = 0$. In other words, all parties mutually benefit without the performance clause in the first-best case. If $v_i > 0$, a risk-averse supplier demands a premium due to the possible penalty associated with the stochastic realization of backorders. This leads to income fluctuations to a risk-averse customer. Both concerns disappear when $v_i = 0$ without incurring extra cost because the contractibility of the suppliers' actions $\{s_i\}$ implies that the actions can be perfectly enforced even without performance incentives. Thus, the customer's attitude toward cost uncertainty and performance uncertainty differ. This key observation will continue to hold even when the suppliers' actions are unobservable and hence not contractible.

4.2 Private Actions: The Suppliers' Problem

We now turn to the situation in which suppliers' actions are unobservable to the customer – which is to be expected in a PBL environment. Given the contract parameters (w_i, α_i, v_i), supplier i chooses (a_i, s_i) that maximize his expected utility (4). That is, he solves

$$\max_{a_i, s_i} w_i - (1-\alpha_i)(c_i s_i - a_i) - v_i E\left[B_i \mid s_i\right] - k_i a_i^2/2 - r_i(1-\alpha_i)^2\text{Var}[\varepsilon_i]/2 - r_i v_i^2 \text{Var}[B_i \mid s_i]/2.$$

A distinctive feature of this problem is that $\text{Var}[B_i \mid s_i]$ is a function of the decision variable s_i. This is a departure from the common assumption found in most moral hazard models that only the mean of the performance measure is affected by the decision variable. In our model the dependence of $\text{Var}[B_i \mid s_i]$ on s_i is unavoidable. As will become evident, this feature of our model complicates the analysis significantly but at the same time creates new dynamics. The supplier's problem is generally not quasiconcave in s_i, but unimodality can be guaranteed under a mild parametric assumption.

Proposition 2 *Suppose $\alpha_i < 1$ and $v_i[1 - F(0)] \geq (1 - \alpha_i)c_i$. Then there is a unique interior optimal solution to the supplier's problem in which supplier i chooses a_i^* and s_i^* such that*

$$a_i^* = (1 - \alpha_i)/k_i, \tag{10}$$

$$v_i[1 - F_i(s_i^*)] + r_i v_i^2 F_i(s_i^*) E[B_i \mid s_i^*] = (1 - \alpha_i)c_i. \tag{11}$$

The condition we specify in Proposition 2 ensures that the supplier's utility function is increasing at $s_i = 0$. The condition has to be checked against the optimal solutions of α_i and v_i, which are determined by the customer (these solutions are presented in the next subsection). We have verified through numerical examples that the condition is mild in the sense that it is violated only under extreme parameter settings (e.g., when the customer's risk aversion factor r_0 is orders of magnitude greater than that of the supplier, r_i). We henceforth assume that the condition in Proposition 2 is always satisfied. From the Proposition we obtain the following results, which offer insights into the impact of contract parameters on optimal decisions.

Corollary 1 *Suppose the conditions in Proposition 2 hold. Then*

(i) $\partial s_i^*/\partial r_i > 0$, $\partial a_i^*/\partial r_i = 0$.

(ii) $\partial s_i^*/\partial \alpha_i > 0$, $\partial a_i^*/\partial \alpha_i < 0$.

(iii) $\partial s_i^*/\partial v_i > 0$, $\partial a_i^*/\partial v_i = 0$.

From (i) we see that the more risk-averse the supplier, the greater the optimal inventory position he chooses. By investing in more spares, the supplier cuts down the likelihood of backorder occurrences, thereby reducing the variance associated with performance. Hence, a risk-averse supplier is inclined to increase s_i to protect himself from performance uncertainty. To put it another way, increasing s_i is a preventive measure that can be taken by the supplier to avoid performance risk. A similar mechanism, however, does not exist for avoiding cost risk, as evidenced by the fact that the optimal cost reduction effort a_i^* is independent of the degree of risk aversion r_i (see equation (10)).[6]

[6] This result is due to the assumption that the stochastic term ε_i enters additively into the supplier's total cost $C_i = c_i s_i - a_i + \varepsilon_i$; the effort reduces the mean of C_i but not the variance. Under this standard assumption the supplier

Parts (ii) and (iii) of Corollary 1 explain optimal supplier responses to the contract terms α_i and v_i. If the customer increases the reimbursement ratio α_i, the supplier becomes less concerned with cost overruns and hence does not exert as much cost reduction effort as he might otherwise ($\partial a_i^*/\partial \alpha_i < 0$). At the same time, his perceived *effective* unit cost of inventory ($(1-\alpha_i)c_i$ on the right-hand side of (11)) decreases, making it desirable to stock more. With respect to the backorder penalty v_i, a larger v_i means a stronger incentive to decrease backorders so s_i^* increases. However, v_i does not affect a_i^*, as it serves only as an incentive to reduce backorders and not the total cost. This behavior is, in part, a consequence of our modeling assumptions that the supplier's effort a_i affects only the cost, and that the uncertainties in cost and in performance are independent of each other.

4.3 Private Actions: The Prime's Problem

Anticipating that the suppliers will respond by choosing $\{a_i, s_i\}$ according to (10) and (11), the customer selects contract terms $\{w_i, \alpha_i, v_i\}$ that achieve minimal total disutility subject to the backorder constraint. With the right incentives, each supplier will voluntarily choose (a_i, s_i) that match the customer's expectation, even though the suppliers' decisions are not directly verified. This voluntary action is expressed in terms of incentive compatibility (IC) constraints, which are added to the customer's problem formulation as follows.

$$
\begin{aligned}
(\mathcal{A}_{SB}) \qquad \min_{\{w_i,\alpha_i,v_i\}} \quad & E\left[U_0\left(\textstyle\sum_{i=1}^n T_i(C_i,B_i)\right) \mid \{a_i^*, s_i^*\}\right], \\
\text{s.t.} \quad & \textstyle\sum_{i=1}^n E\left[B_i \mid s_i^*\right] \le \widehat{B}_0, && \text{(AR)} \\
& E\left[U_i\left(T_i(C_i,B_i) - C_i\right) - \psi_i(a_i) \mid a_i^*, s_i^*\right] \ge 0 \qquad \forall i, && (\text{IR}_i) \\
& (a_i^*, s_i^*) \in \arg\max E\left[U_i\left(T_i(C_i,B_i) - C_i - \psi_i(a_i)\right) \mid a_i, s_i \ge 0\right] \quad \forall i, && (\text{IC}_i) \\
& 0 \le \alpha_i \le 1 \qquad \forall i.
\end{aligned}
$$

Similar to the first-best case, it can be demonstrated that the (IR_i) constraints bind at the optimum, so that we can simplify the problem by solving for values of $\{w_i\}$ that leave the suppliers with zero expected utility. Using the Lagrange multiplier θ for the backorder constraint, we can write n individual Lagrangian functions. Moreover, it is convenient to convert the Lagrangian into a function of (α_i, s_i, θ) rather than a function of (α_i, v_i, θ), using the monotonicity result $\partial s_i^*/\partial v_i > 0$ from Corollary 1. Using

has no control over the variability of cost, so his attitude toward risk does not factor into the decision about a_i^*.

(10), we obtain

$$\begin{aligned}\mathcal{L}_i(\alpha_i, s_i, \theta) \;&=\; c_i s_i + \theta E\left[B_i \,|\, s_i\right] - (1-\alpha_i)/k_i + (1-\alpha_i)^2/\left(2k_i\right) + \left(r_0 {\alpha_i}^2 + r_i(1-\alpha_i)^2\right) Var[\varepsilon_i]/2 \\ &\quad + (r_0 + r_i)\left[v_i(\alpha_i, s_i)\right]^2 \mathrm{Var}[B_i \,|\, s_i]/2, \end{aligned} \tag{12}$$

whereby

$$v_i(\alpha_i, s_i) = \begin{cases} \frac{(1-\alpha_i)c_i}{1-F_i(s_i)} & \text{if} \quad r_i = 0, \\ \frac{1-F_i(s_i)}{2r_i F_i(s_i) E[B_i \,|\, s_i]}\left(-1 + \sqrt{1 + \frac{4r_i c_i (1-\alpha_i) F_i(s_i) E[B_i \,|\, s_i]}{[1-F_i(s_i)]^2}}\right) & \text{if} \quad r_i > 0, \end{cases} \tag{13}$$

from (11). We readily notice that the optimal performance incentive $v_i(\alpha_i, s_i)$ is a decreasing function of α_i; in order to have the supplier choose s_i, the customer may decrease v_i while increasing α_i, or vice versa. Thus, v_i, the incentive to increase the stocking level, and $1-\alpha_i$, the incentive to reduce costs, are complements. This observation plays a key role in a later analysis and will be discussed further.

We denote the optimal solution pairs with superscripts SB, $\{\alpha_i^{SB}, s_i^{SB}\}$. Unfortunately, (12) is not generally quasiconvex and hence is not necessarily unimodal. The analytical specification of s_i^{SB} is intractable even with α_i fixed, thereby requiring numerical analysis. To gain additional insights while circumventing this difficulty, in the next section we focus on several special cases and later analyze the original problem numerically.

4.4 Cost Plus (C+) vs. Fixed Price (FP) vs. Performance Contracts

Before delving into the analysis of optimal contracts for special cases, we pause here to evaluate the effectiveness of the most widely used contract forms, C+ ($\alpha_i = 1$, $v_i = 0$) and FP ($\alpha_i = v_i = 0$), and compare them with performance contracts ($v_i > 0$). Consistent with other literature analyzing and comparing these contracts (see, for example, Scherer [27]), our model indicates that C+ and FP are polar opposites when it comes to providing cost reduction incentives. With a FP contract a supplier becomes the residual claimant and hence it is in his interest to reduce costs as much as possible. In terms of cost risk, the FP contract gives perfect insurance to the customer because the supplier bears all risks from cost under- or overruns. In contrast, the C+ contract shifts all risks to the customer, as she has to reimburse whatever the supplier's realized cost may be. At the same time, the C+ contract provides no incentive for the supplier to reduce costs.

Despite the prevalence of C+ and FP contracts in practice, they do not induce the desired supplier behavior when a performance constraint is present and the customer cannot observe suppliers' actions. This becomes clear after inspecting the supplier's utility function (4). With the FP contract, it is in

Contract type	No performance-based compensation ($v = 0$)	Performance-based compensation ($v > 0$)
Pure performance ($\alpha = w = 0$)		The customer is unable to extract all supplier surplus.
Fixed price ($\alpha = 0$)	While achieving the first-best cost reduction effort a^{FB}, the supplier is incentivized to reduce s as much as possible.	First-best can be achieved with the appropriate choice of w and v under risk neutrality. First-best is not achieved under risk aversion ($\alpha > 0$) in general.
Cost plus ($\alpha = 1$)	The supplier exerts no cost reduction effort ($a = 0$) and is indifferent toward s.	The supplier exerts no cost reduction effort ($a = 0$) and tries to increase s as much as possible.

Table 1: Incentive effects of various contract combinations.

the supplier's interest to reduce not only effort a_i but also spares inventory s_i as much as possible, thus violating the minimum availability desired by the customer. A C+ contract, on the other hand, has the effect of making the supplier indifferent to the choice of s_i. Clearly, inducing proper actions requires performance incentives. The simplest contract in this category (the "pure performance contract") has $\alpha_i = w_i = 0$ and $v_i > 0$. Indeed, such a contract can induce the supplier to choose the optimal inventory level s_i. But it is inefficient because it leaves a positive residual surplus to the suppliers (i.e., the (IR) constraint does not bind in general). Interestingly (to be demonstrated in the following section) a contract with $w_i > 0$ and $v_i > 0$ can achieve the first-best solution. But this happens only if all parties are risk-neutral, as proper risk sharing requires $\alpha_i > 0$. Thus, the optimal contract will have all three components: a fixed payment, a cost-sharing clause, and a performance incentive. Table 1 summarizes supplier behavior under all of these contract combinations.

5 Special Cases

5.1 Risk-Neutral Firms

Many difficulties associated with the analysis disappear if all suppliers and the customer are risk-neutral, which may be the case in practice if the customer and the suppliers are all very large, well-diversified corporations. In this case, as we show below, even when actions are unobservable, the first-best solution is achieved with a contract that is a simple combination of a fixed payment and a performance component (henceforth called FP/performance). This solution highlights the performance allocation aspect of our problem at the expense of ignoring the issue of risk-sharing.

Proposition 3 *With $r_0 = r_1 = ... = r_n = 0$, the first-best solution is achieved if and only if*

(i) $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$,

(ii) $w_i = c_i s_i^{FB} + \theta^{FB} E[B_i \,|\, s_i^{FB}] - 1/2k_i$,

(iii) $v_1 = v_2 = ... = v_n = \theta^{FB}$

where $\{s_i^{FB}\}$ *and* θ^{FB} *are computed from (7) and (8). The supplier i's expected utility is zero while the customer's expected utility is* $\sum_{i=1}^{n} \left(-c_i s_i^{FB} + 1/2k_i\right)$.

The preceding result is not entirely new: it is often the case in other principal-agent models that the first-best solution is achieved with an FP/performance contract between two risk-neutral firms when there is only one effort variable (for example, see Bolton and Dewatripont [2]). Having two effort variables a_i and s_i as well as multiple suppliers does not change this basic result. First-best is obtained because α_i and v_i under risk neutrality serve only as incentives and not as instruments for providing insurance against risk, eliminating the trade-off between the two factors.

There is, however, an interesting deviation from the classical analysis involving just one supplier. It is captured in part (iii), which can be interpreted to mean that every backorder from heterogeneous subsystems has equal importance regardless of the subsystem unit price c_i. Thus, performance incentives are equal across suppliers. In our additively separable backorder model ($B_0 = \sum_{i=0}^{n} B_i$) this makes intuitive sense, because the customer does not discriminate between a backorder of a \$1,000 item and that of a \$10 item; hence each item contributes equally to the downtime of the system. However, it would be erroneous to conclude that item unit costs $\{c_i\}$ have no effect on determining the uniform performance incentive $v_1 = ... = v_n = \theta^{FB}$ because they determine θ^{FB} indirectly through the joint satisfaction of (7) and (8). The fact that penalty rates are linked across suppliers continues to hold in the risk-averse case, although the equality as in (iii) can no longer be sustained because of the suppliers' varying attitudes toward risk. The policy implication of this result is to treat all suppliers equally with respect to the performance incentive as long as risk aversion is not present.

5.2 Risk-Averse Firms: Cases with Partial Observability

As the next step in gaining insights we now analyze the problem under a simplifying assumption that either $\{s_i\}$ or $\{a_i\}$ are observable and contractible, but not both. As will become evident, these special cases serve as bounds on the optimal contract parameters under conditions of complete unobservability and hence are useful in understanding the structure of the problem. We shall first consider the case when $\{s_i\}$ are observable but $\{a_i\}$ are not. This may happen if the suppliers utilize consignment inventory management for all subsystems (which is sometimes the case in practice) so that inventories are visible to the customer. As s_i can now be dictated by the customer, there is no need to provide the performance incentive v_i, i.e., the optimal contract has $v_i = 0$ for all i. After determining $\{w_i\}$ from the binding (IR) constraints, the customer's problem ($\mathcal{A}'_{SB}$) becomes

$$(\mathcal{A}'_{SO}) \quad \min_{\{\alpha_i, s_i\}} \quad \sum_{i=1}^n \left(c_i s_i - (1-\alpha_i)/k_i + (1-\alpha_i)^2 / (2k_i) + \left(r_0 \alpha_i^2 + r_i (1-\alpha_i)^2 \right) \mathrm{Var}[\varepsilon_i]/2 \right),$$
$$\text{s.t.} \quad \sum_{i=1}^n E\left[B_i \,|\, s_i\right] \le \widehat{B}_0.$$

The optimal contract (denoted by the superscript SO) is as follows.

Proposition 4 *When $\{s_i\}$ of all suppliers are observable to the customer but $\{a_i\}$ are not, it is optimal to specify the contract terms according to*

(i) $\alpha_i^{SO} = k_i r_i / \left(1/Var[\varepsilon_i] + k_i(r_0 + r_i)\right) < \alpha_i^{FB}$,

(ii) $w_i^{SO} = (1-\alpha_i^{SO}) c_i s_i^{FB} - (1-\alpha_i^{SO})^2 / (2k_i) + r_i (1-\alpha_i^{SO})^2 \, Var[\varepsilon_i]/2$, *and*

(iii) $v_i^{SO} = 0$.

$s_i^{SO} = s_i^{FB}$ *is imposed on supplier i while the contract terms induce the cost reduction effort* $a_i^{SO} = (1 + k_i r_0 Var[\varepsilon_i]) / \left(k_i + k_i^2 (r_0 + r_i) Var[\varepsilon_i]\right)$.

Even though one of the supplier's actions is observable to the customer, we see that the first-best solution cannot be achieved and hence there are inefficiencies due to incentive issues. Namely, there is less cost sharing than is optimal under the first-best solution ($\alpha_i^{SO} < \alpha_i^{FB}$) because the customer has to provide more incentive to reduce costs than would have been the case if she dictated a_i, and this is achieved by exposing the supplier i to more risk (smaller α_i). We see that α_i^{SO} exhibits intuitive properties: as $\mathrm{Var}[\varepsilon_i]$ approaches infinity, α_i^{SO} increases asymptotically to the first-best optimal risk sharing ratio α_i^{FB} because the supplier's effort a_i becomes overshadowed by huge cost uncertainty. It is also clear that α_i^{SO} moves toward zero (toward an FP contract) as $\mathrm{Var}[\varepsilon_i]$ decreases. The relative risk aversion ratio r_0/r_i is another major determinant of α_i^{SO} which is similar to the first-best case: if the ratio is small, α_i^{SO} is on the C+ side (closer to 1) whereas a large ratio implies that α_i^{SO} is on the FP side (closer to 0).

The other possibility is when $\{a_i\}$ of all suppliers are observable but $\{s_i\}$ are not. This situation could arise in government contracting where a significant amount of information on supplier costs must be divulged to the customer.[7] We denote the optimal solution in this case with the superscript AO. As before, $\{w_i\}$ are determined from the binding (IR) constraints. The customer's problem becomes

$$(\mathcal{A}'_{AO}) \quad \min_{\{a_i, \alpha_i, s_i\}} \quad \sum_{i=1}^n \Big(c_i s_i - a_i + k_i a_i^2/2 + \left(r_0 \alpha_i^2 + r_i(1-\alpha_i)^2 \right) Var[\varepsilon_i]/2 + (r_0 + r_i) \left[v_i(\alpha_i, s_i) \right]^2 \mathrm{Var}[B_i \,|\, s_i]/2 \Big)$$
$$\text{s.t.} \quad \sum_{i=1}^n E\left[B_i \,|\, s_i\right] \le \widehat{B}_0$$

[7] The Truth in Negotiations Act (TINA) has been applied to many government contracts since the 1960's. It requires suppliers to reveal cost data to the government (customer) in order to avoid excessive payments to the suppliers. In most PBL contracts, however, TINA is waived.

It is clear that $a_i^{AO} = a_i^{FB}$ as in (6) but tractable expressions for α_i^{AO} and s_i^{AO} do not exist. Despite this shortcoming, α_i^{AO} can be evaluated analytically in the special case with only one supplier, a scenario which we present next.

5.3 Single Risk-Averse Supplier

In this subsection we assume that there is only one supplier, so we drop the subscript i. Not only is such a firm-to-firm setting consistent with a majority of supply chain contracting models in the literature, but it is also a commonly observed situation that is found in PBL practice. For example, a setting in which maintenance of a single key component is outsourced or a military customer contracts directly with a subsystem supplier fits this description (e.g., the U.S. Navy's PBL contract with Michelin for tires or commercial airline "power by the hour" contracts with engine manufacturers like GE and Rolls Royce). As we will see shortly through numerical experiments, insights from this simpler model continue to hold for the general assembly structure with multiple suppliers.

With a single supplier, it may appear that the customer should set incentives so that $E\left[B \mid s^{SB}\right] = \widehat{B}_0$ holds. In particular, this would be the case if the customer's objective funciton were increasing monotonically in s, which is an intuitive property. Unfortunately, this intuition is not entirely correct. As noted in the previous section, the analysis of risk-averse firms is complicated by the non-quasiconcavity, implying that the Lagrangian (12) can be bimodal. Thus, the customer may prefer to have more inventory than follows from $E\left[B \mid s^{SB}\right] = \widehat{B}_0$. This, however, happens only in extreme cases in which the customer is several orders of magnitude more risk-averse than the supplier and therefore wants to protect herself from performance risk with a very large inventory. In most of our numerical examples, which cover a wide range of parameter combinations, the customer's objective function is, indeed, increasing monotonically in s. Therefore, we will henceforth assume that the problem parameters are such that the backorder constraint is binding, so that the optimal inventory position s^{SB} satisfies $E\left[B \mid s^{SB}\right] = \widehat{B}_0$. Given that v is completely determined by α and s according to (13), the only variable to be determined is the cost-sharing parameter α so that our problem is simplified to a one-dimensional optimization.

Lemma 1 *The customer's Lagrangian (12) is convex in α when s is fixed.*

It follows that there is a unique α^{SB} that minimizes the customer's objective function. There exists a closed-form solution, but it is quite complex (see the Appendix) and inspection alone does not provide ready insights. Instead, we focus on understanding how the parameters of the contract change when cost uncertainty $\text{Var}[\varepsilon]$ changes. There are several motivations behind this analysis. First, cost uncertainty is of primary importance in practice because it is often harder to estimate than performance uncertainty. Second, there are significant changes in cost uncertainty over the product life cycle (while performance uncertainty can be relatively more stable) and therefore there is a need to understand how contractual terms should change in response. Finally, as will be seen shortly, by varying cost uncertainty we are able to obtain insights that sometimes differ fundamentally from insights in the classical literature on moral hazard problems with multitasking.

Proposition 5 *Suppose $r_0, r > 0$ and that s^{SB} is fixed by the backorder constraint $E\left[B \mid s^{SB}\right] = \widehat{B}_0$. Then $\alpha^{SO} < \alpha^{SB} < \alpha^{AO}$ and $v^{SB} > v^{AO} > v^{SO} = 0$. Further, let $\ell(r_0, r) = \partial\mathcal{L}/\partial\alpha|_{\alpha=\alpha^{FB}}$ where $\mathcal{L}$ is the customer's Lagrangian defined in (12). Function $\ell(r_0, r)$ increases in the ratio r/r_0 and crosses zero exactly once. The optimal contract parameters α^{SB} and v^{SB} are related to α^{FB} and v^{FB} as follows.*

(i) If $\ell(r_0, r) > 0$, $\alpha^{SB} < \alpha^{FB}$, $d\alpha^{SB}/d(Var[\varepsilon]) > 0$, and $dv^{SB}/d(Var[\varepsilon]) < 0$.

(ii) If $\ell(r_0, r) = 0$, $\alpha^{SB} = \alpha^{FB}$, $v^{SB} = v^{FB}$, and $d\alpha^{SB}/d(Var[\varepsilon]) = dv^{SB}/d(Var[\varepsilon]) = 0$.

(iii) If $\ell(r_0, r) < 0$, $\alpha^{SB} > \alpha^{FB}$, $d\alpha^{SB}/d(Var[\varepsilon]) < 0$, and $dv^{SB}/d(Var[\varepsilon]) > 0$.

First, we note that the optimal cost sharing ratio α^{SB} is bounded above by α^{AO}, the optimal ratio when the cost reduction effort a is observable. In the current case the effort is not observable and therefore the customer has to reduce α to provide more incentives to reduce costs. The side effect is that the supplier's effective unit cost $(1-\alpha)c$ increases, thus requiring a higher performance incentive v to induce the desired inventory position s^{SB}. Therefore, $v^{SB} > v^{AO}$. Second, we note that α^{SB} is bounded below by α^{SO}, which we derived by assuming that the inventory position s is observable. When s is not observable, the customer needs to provide a higher performance incentive, $v^{SB} > v^{SO} = 0$. But doing so exposes both the customer and the supplier to performance risk (recall the performance risk premium is increasing in v for both the customer and the supplier; see (4) and (5)), thus creating inefficiency that can be mitigated by increasing α. Higher α reduces the effective unit cost $(1-\alpha)c$ for the supplier and allows him to achieve the inventory position s^{SB} with a smaller v. Hence, increasing α above α^{SO} is optimal.

A comparison of the second-best solution with the first-best solution is more complex. It is instrumental to consider two cases based on the relative risk aversion of the customer and the supplier separately. Since function $\ell(r_0, r)$ increases in the ratio r/r_0 and crosses zero exactly once, the condition $\ell(r_0, r) > 0$ in (i) can be interpreted as $r \succ r_0$, where the symbol "$\succ$" means that the supplier is relatively more risk-averse than the customer. Similarly, $\ell(r_0, r) < 0$ can be interpreted as $r \prec r_0$, whereby the customer is relatively more risk-averse than the supplier. We first consider the former situation (which may arise if the customer is a bigger and more diversified company than the supplier). We believe that this case is more natural in practice. Figure 2 illustrates the results in (i).

We make the following observations from these figures. First, $\alpha^{SB} < \alpha^{FB}$, and the unobservability

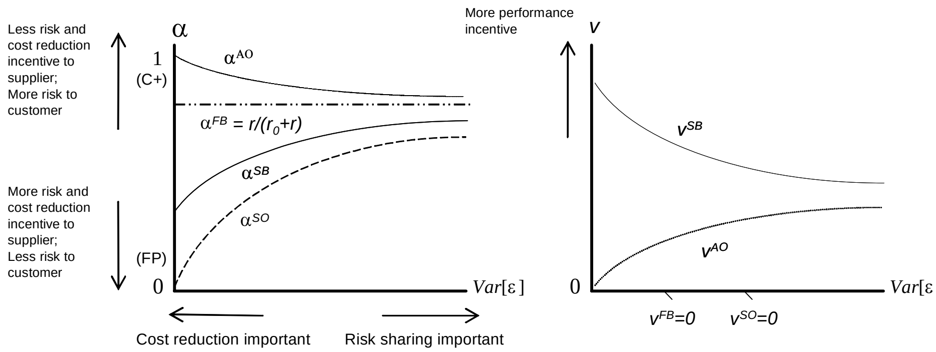

Figure 2: $\ell(r_0, r) > 0$, the supplier is relatively more risk-averse than the prime.

of effort and inventory results in less cost reimbursement than under the first-best solution. Second, α^{SB} increases with Var$[\varepsilon]$ and asymptotically approaches α^{FB}. With large cost uncertainty, the risk-averse supplier is reluctant to participate in the trade, so the customer has to provide insurance by reimbursing a large proportion of the supplier's costs. Thus the supplier has less incentive to make efforts to reduce costs. On the other hand, when Var$[\varepsilon]$ is small, providing cost-reduction incentives becomes more important. Third, the gap between α^{SB} and α^{SO} decreases in Var$[\varepsilon]$. This gap can be interpreted as the additional inefficiency attributed to performance risk. When cost uncertainty is large, performance uncertainty Var$[B \,|\, s^{SB}]$ is negligible and the gap between SB and SO disappears. The gap between α^{SB} and α^{AO} is interpreted similarly. Finally, v^{SB} decreases with Var$[\varepsilon]$, asymptotically approaching $v(\alpha^{FB}, s^{FB})$. With higher cost uncertainty, the performance incentive is lowered.

Overall, we observe that α^{SB} and v^{SB} move in the opposite directions as Var$[\varepsilon]$ increases because the customer increases α to mitigate the supplier's cost risk (we recall that the supplier is more risk-averse than the customer in the current setting). As a result, the supplier's effective unit cost $(1-\alpha)c$ is reduced, making it less expensive to stock inventory and allowing for a smaller incentive v. Therefore increasing $1-\alpha$ has the same effect on inventory as increasing v; these two incentives are *complements* with respect to s.

This conclusion is similar to the one presented in Holmström and Milgrom's [14] original multitask principal-agent model in which increasing variability in one output leads to weaker incentives for all outputs. Yet, the mechanism by which we arrive at our conclusion is different. Specifically, in Holmström and Milgrom [14], raising one effort raises the marginal disutility of raising another effort, which is not the case in our model (since the supplier's disutilities $(1-\alpha)cs$ and $ka^2/2$ are independent of each other). Another important assumption in their model is that the outcomes are affected by exactly one effort each, so there is a one-to-one correspondence between an incentive and an effort. In contrast, our model has an outcome C that is a function of both variables a and s via $C = cs - a + \varepsilon$.

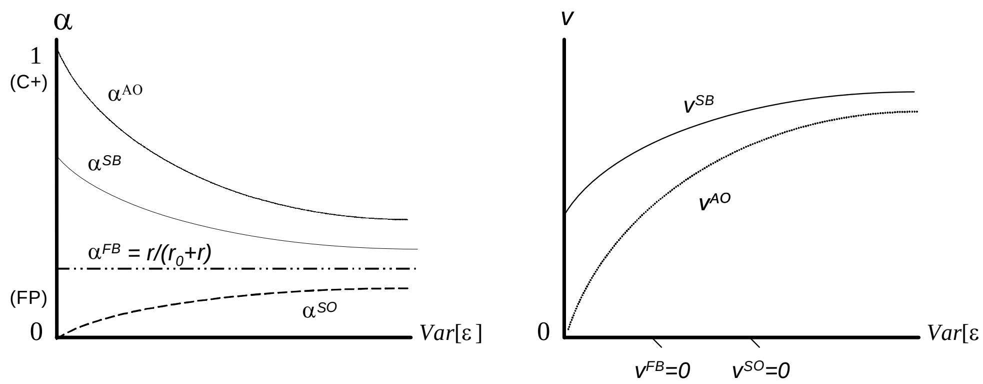

Figure 3: $\ell(r_0, r) < 0$, the prime is relatively more risk-averse than the supplier.

In this respect, the model closest to ours is found in Bolton and Dewatripont ([2], pp. 223-8) where there is *direct conflict between the tasks,* because exerting one effort positively affects one outcome but negatively affects the other.

Next, we consider the case in which the customer is relatively more risk-averse than the supplier, $r \prec r_0$ (case (iii) in Proposition 5). Figure 3 is an analog of Figure 2. Compared to the previous discussion, α^{SB} and v^{SB} exhibit exactly opposite behavior. Now $\alpha^{SB} > \alpha^{FB}$ and α^{SB} decreases in $\text{Var}[\varepsilon]$ while v^{SB} increases in $\text{Var}[\varepsilon]$. This fundamental difference arises because, unlike in the previous case, it is now the customer who needs more protection from cost risk. In the presence of large cost uncertainty, this can be achieved by choosing small α, thereby transferring most of the risk to the supplier. A nonintuitive consequence of this outcome is that the supplier is incentivized more to reduce his cost and increase his stocking level when cost uncertainty is great. Therefore, the customer's concern for her own risk protection reverses contractual terms and comparative statics. The complementarity between $1 - \alpha^{SB}$ and v^{SB} still remains, however: as $1 - \alpha^{SB}$ increases, so does v^{SB}. We note that results when the customer is more risk-averse than the supplier are somewhat contrary to what we have come to expect from the existing literature on multitasking where the customer is often assumed to be risk-neutral.

6 Examples with Multiple Risk-Averse Suppliers

In this section we present a numerical analysis of the problem with multiple suppliers. We illustrate our findings through two examples. First, we consider two suppliers that differ by at most one of the parameters $\{r_i, \text{Var}[\varepsilon_i]\}$. This example isolates the trade-off between incentives and risk. The second example is based on actual maintenance data from a fleet of military fighter aircraft. This second data set illustrates how our model can be applied in practice to support long-term strategic planning and contract negotiations.

6.1 Example 1: Two Symmetric Suppliers

In this example we assume that all parameter values are symmetric across the suppliers except for either $\{r_i\}$ or $\{\text{Var}[\varepsilon_i]\}$. Default values are $\mu_i = \sigma_i^2 = 10$, $c_i = 1$, $k_i = 0.2$, $r_1 = 0.1$, $\text{Var}[\varepsilon_1] = 10$, and $\widehat{B}_0 = 4$. A normal distribution of the inventory on-order is chosen, in keeping with our continuous approximation for the underlying inventory model. We vary supplier 2's risk aversion r_2 and cost uncertainty $\text{Var}[\varepsilon_2]$ as well as the customer's risk aversion r_0 in order to observe their effects on $\{\alpha_i^{SB}, v_i^{SB}\}$ and $\{a_i^{SB}, s_i^{SB}\}$. We note that the first-best inventory positions are $s_1^{FB} = s_2^{FB} = 8.725$. Table 4 (see Appendix) summarizes the results of varying r_2 and r_0.

We observe only minimal changes in s_1^{SB} and s_2^{SB} as parameters change, with the greatest change occurring when r_0 is large. In contrast, α_2^{SB} changes widely and so does v_2^{SB}, but in the opposite direction. As we observed in the single supplier case earlier, the complementarity between the cost reduction incentive $1 - \alpha_i$ and the performance incentive v_i continues to exist in the multiple supplier setting. $\{s_i^{SB}\}$ do not vary much because they are subject to an externality, namely, the overall backorder constraint. Although $\{s_i^{SB}\}$ can be used as instruments for protection from performance risk, their ranges are limited by the constraint. Hence, risk allocation (including allocation of the performance risk) is primarily achieved through varying $\{\alpha_i\}$.

From the table we also confirm that $\alpha_2^{SB} > \alpha_2^{FB}$ for a relatively large ratio r_0/r_2, while the opposite is true for a small ratio r_0/r_2, just as predicted by Proposition 5 but for a single supplier. Furthermore, we notice that α_2^{SB} increases monotonically in r_2 for small r_0 ($= 0.01$), but we do not observe the same monotonicity when r_0 is large ($= 1$): α_2^{SB} initially decreases from 0.436 to 0.430 but then increases to 0.539. The explanation is as follows. When r_0 is small, increasing α_i tends to reduce both the cost and performance premiums (see (12) and (13)). However, when r_0 is large, tension exists between the two risk premium terms; although increasing α_i reduces the performance risk for both the customer and the suppliers, it exposes the customer to the risk of greater cost. These two opposing forces break down the monotonicity.

Next, Table 5 (see Appendix) illustrates the effect of varying $\text{Var}[\varepsilon_2]$. Once again we observe that s_2^{SB} is not very sensitive to changes in $\text{Var}[\varepsilon_2]$, but that α_2^{SB} is. From the table, we see that α_2^{SB} moves toward α_2^{FB} as $\text{Var}[\varepsilon_2]$ increases (regardless of the value of r_0), confirming the prediction from Proposition 5 for the single supplier case. In addition, numbers in the table indicate that $\partial\alpha_2^{SB}/\partial r_0 > 0$ for small $\text{Var}[\varepsilon_2]$ whereas $\partial\alpha_2^{SB}/\partial r_0 < 0$ for large $\text{Var}[\varepsilon_2]$ (this can be proven analytically in the single supplier case, but we omit the derivation). In other words, when cost uncertainty is relatively low, a more risk-averse customer moves toward a C+ contract and takes up a larger portion of the cost risk, a nonintuitive result. What actually happens is that the performance premium is more important in this situation, outweighing the concern for the cost risk. Clearly, in the presence of performance risk, the intuition regarding cost sharing is not always straightforward.

Finally, Table 6 (see Appendix) shows how the optimal contract parameters and the suppliers' actions vary as the overall backorder constraint changes. In this example, suppliers 1 and 2 are asymmetric only in their attitude toward risk: $r_1 = 0.1$ and $r_2 = 1$, while $r_0 = 0.5$. As expected, v_1^{SB} and v_2^{SB} decrease as $\widehat{B}_0$ increases, since a less stringent backorder constraint allows for smaller inventories and hence reduces the need for performance incentives. Changes in α_1^{SB} and α_2^{SB} are relatively small. We see that distortion in $\{s_i^{SB}\}$ becomes larger as we relax the constraint (measured by the quantity $(s_2^{SB} - s_1^{SB})/s_1^{FB}$, it grows from 2.94% at $\widehat{B}_0 = 1$ to 9.75% at $\widehat{B}_0 = 7$). Intuitively, this happens because the less stringent backorder constraint results in a larger range in which inventories can be adjusted without violating the constraint. However, the magnitude of the distortion is still small, confirming our previous observation that the presence of the backorder constraint limits the customer's contract parameter choices (α and v) such that they induce the base stock levels close to the first-best values $\{s_i^{FB}\}$.

6.2 Example 2: Actual Data for a Fleet of Military Aircraft

Our second numerical example is based on a real-life maintenance data for a fleet of military fighter aircraft. A total of $N = 156$ aircraft are deployed in the fleet. We obtained data on unit costs, daily failure rates and repair lead times for a representative collection of 45 line replaceable units ("parts"). To utilize our model we aggregate data into five subsystem groups: avionics (a), engines (e), landing gear (l), mechanical (m), and weapons (w), based on descriptions of each part. We employ the following technique to obtain unit costs, failure rates and lead times for these subsystems. First, we assign each part to one of the groups, and compute the subsystem's mean inventory on-order as $\mu_i = \sum_{j=1}^{n_i} \lambda_j L_j$, where i ={a,e,l,m,w} and n_i = the number of parts within subsystem i. Thus, we treat each subsystem as a "kit" which is replaced whenever any part within it fails. Subsystem unit costs are inferred from an output generated by the proprietary commercial software from MCA Solutions, Inc. (http://www.mcasolutions.com). Given a system availability target, this software calculates optimal stocking levels over multiple echelons and indentures while directly considering each part-location. By aggregating its output, we can infer the effective subsystem unit costs by dividing the dollar amount invested in inventory resources for each subsystem ($\sum_{j=1}^{n_i} c_j s_j$) by the total number of stocking units within it ($\sum_{j=1}^{n_i} s_j$). For this example the availability target of 95% was chosen. Table 2 summarizes the inferred values of $\{\mu_i\}$ and $\{c_i\}$ using this heuristic. We note that $\{\mu_i\}$ are an order of magnitude smaller than N, thus satisfying the condition $E[B_i \,|\, s_i] \ll N$ needed to apply the fixed failure rate

Subsystem	avionics (a)	engine (e)	landing gear (l)	mechanical (m)	weapons (w)
μ_i	10.46	19.36	13.72	16.87	8.43
c_i (in \$1,000)	21.52	6.60	31.08	8.52	14.85

Table 2: μ_i and c_i for each subsystem.

	$\rho_i = 0.02$					$\rho_i = 0.1$				
i	a	e	l	m	w	a	e	l	m	w
α_i^{SB}	0.54	0.67	0.88	0.41	0.51	0.63	0.35	0.92	0.48	0.75
v_i^{SB}	5.77	8.11	1.46	6.49	4.35	5.01	12.61	1.14	5.97	3.06
a_i^{SB}	99.74	21.66	38.15	50.58	73.10	79.14	43.11	25.01	44.04	37.52
s_i^{SB}	9.29	23.12	10.51	19.26	8.66	9.35	22.77	10.39	19.27	8.96
$\widehat{A}_i$	98.75%	99.69%	97.69%	99.54%	99.33%	98.77%	99.65%	97.63%	99.54%	99.42%
IIR_i	200.0	152.6	326.8	164.1	128.6	201.2	150.3	322.9	164.2	133.1
NCR_i	76.6	18.1	35.8	35.6	55.1	64.6	29.0	24.0	32.7	32.8
CRP_i	23.4	1.5	68.6	2.8	17.6	557.3	15.1	1669.3	68.1	318.8
PRP_i	35.9	10.0	19.8	15.9	16.1	26.7	28.2	12.2	13.4	7.0

Table 3: Optimal contract terms and suppliers' actions. The dollar figures are in thousands. IIR stands for investment in resources and is equal to $c_i s_i^{SB}$. NCR is $-a_i^{SB} + \frac{1}{2}k_i(a_i^{SB})^2$, the net cost reduction. CRP is the residual cost risk premium, $\frac{1}{2}(r_0(\alpha_i^{SB})^2 + r_i(1-\alpha_i^{SB})^2)\text{Var}[\varepsilon_i]$, and PRP is the residual performance risk premium, $\frac{1}{2}(r_0 + r_i)(v_i^{SB})^2\text{Var}[B_i|s_i^{SB}]$. System availability target is 95%.

approximation.

To determine values of parameters $\{k_i\}$ and $\{\text{Var}[\varepsilon_i]\}$, we use the following approach. Let K_i be the supplier i's fixed cost such that $K_i = E[K_i] + \varepsilon_i$. For each supplier, we assume that the expected fixed cost is 50 times higher than the unit cost c_i. The maximum dollar amount of cost reduction $a_i^{FB} = 1/k_i$ is assumed to be $0.2E[K_i]$. Thus, $k_i = 1/(10c_i)$. For the sake of simplicity, we also assume that the coefficient of variation $\rho_i \equiv \sqrt{\text{Var}[K_i]}/E[K_i]$ is the same across suppliers. We infer the risk aversion coefficient for each supplier from the market capitalization of a representative manufacturer of such a subsystem. For example, if Boeing is chosen as the customer and GE as the engine supplier, we calculate the risk aversion ratio of $r_0/r_e \simeq 7$ since GE's market capitalization is roughly 7 times that of Boeing (see justification for using company size as a proxy for risk aversion in Cummins [9]). This approach is, of course, quite simplistic, but it fits our aim to illustrate the model. Using this methodology we choose $r_a/r_0 = 1.79$, $r_e/r_0 = 0.15$, $r_l/r_0 = 11.76$, $r_m/r_0 = 1$, and $r_w/r_0 = 3.33$ and we select $r_0 = 0.15$. The optimal contract terms and the suppliers' actions are presented in Table 3.

We consider two scenarios: with small and high cost uncertainty (as captured by the coefficient of variation ρ_i). For simplicity, assume that all suppliers have the same value of ρ_i. Table 3 summarizes optimal contract parameters and the implied cost terms, including the cost and performance premiums. In the case of high uncertainty, observe that the cost premium is higher than the performance premium for all suppliers except for the engine supplier (e). This asymmetry arises because he is the only supplier who is less risk averse than the customer. On the other hand, the performance premium becomes more salient when cost uncertainty is small. We also observe that $\{\alpha_i^{SB}\}$ increases and $\{v_i^{SB}\}$ decreases with

ρ_i for all suppliers except (e), which is consistent with our results for a single supplier. Finally, the values of $\{s_i^{SB}\}$ are quite insensitive to changes in ρ_i as noted in the previous example.

7 Conclusion

The goal of this paper is to introduce contracting considerations into the management of after-sales service supply chains. We do so by blending the classical problem of managing the inventory of repairable service parts with a multitask principal-agent model. We use this model to analyze incentives provided by three commonly used contracting arrangements, fixed-price, cost-plus and performance-based (FP, C+ and PBL). By doing so, we analyze two practically important issues of contracting in service supply chains – performance requirement allocation and risk sharing – when a single customer is contracting with a collection of first-tier suppliers of the major subsystems used by an end product/system. When performance is defined as overall system availability, the answer to the former can be found from the solution of the classic service part resource allocation problem. Our innovation is in explicitly modeling decentralized decision making and considering how firms behave when they face uncertainties arising from both support costs and product performance. The notion of risk sharing found in the principal-agent literature is incorporated into our model, providing insights into what types of contracts should be used under various operating environments. Specifically, we have discovered that incentive terms in the contract exhibit complementarity, i.e., incentives for both cost reduction and high availability move in the same direction as cost uncertainty changes.

Furthermore, our analysis allows us to make normative predictions with respect to how contracts are likely to evolve over the product life cycle. Given our assumption that supplier effort reduces support costs but does not improve product performance reliability or repair capabilities, our model is consistent with the observation that performance uncertainty is relatively stable throughout the repair and maintenance process, while cost uncertainty is likely to be reduced over time by learning about costs through the deployment of a larger fleet of systems. Thus, if a series of performance contracts are signed over the product lifetime, our analysis indicates that the cost reimbursement ratio α will decrease (increase) over time if the supplier is relatively more (less) risk-averse than the customer. For the performance incentive v the direction is reversed. Since larger, more diversified customers are more common in practice, our results predict that the optimal contract will typically assume less cost sharing and more performance incentive as the product matures. Indeed, this prediction is confirmed by practitioners and from the DoD publications: "PBL strategies will generally have a phased contracting approach, initiated by Cost Plus cost reimbursement type contracts to Cost Plus incentive contracts to Fixed Price incentive contracts, over time" (Defense Acquisition University [12]).

We find that, in the presence of great residual uncertainty associated with performance, cost sharing is still an effective tool even if cost uncertainty is small. That is, the combination FP/performance-based contract is not optimal in such instances (notice the gap between zero and α^{SB} at $\text{Var}[\varepsilon] = 0$ in Figure 2), because the cost reimbursement α can be used as a risk protection mechanism even for the risk borne by the performance. While inventory s can be used as an instrument to hedge against performance risk, adjusting α is more effective for this purpose because the primary role of s is controlling for the backorder level in order to achieve the availability target. Hence, some degree of cost sharing is recommended in a performance contracting environment even when cost uncertainty is low. Our numerical study shows that the optimal inventory position profile $\{s_i^{SB}\}$ is quite insensitive to changes in risk-related parameters such as r_0, r_i, and $\text{Var}[\varepsilon_i]$. This happens because the presence of a stringent backorder constraint limits the range in which $\{s_i^{SB}\}$ can be varied.

Performance-based contracting in service supply chains offers fertile ground for research where economics and classical inventory theory converge naturally. Not only does it pose theoretically challenging questions, but insights gained from the analysis are of great interest to practitioners who are currently undergoing major business process changes due to the move towards PBL contracting. Our paper analyzes several major issues in performance contracting, but many open questions remain. Follow-up studies may address such topics as the free-riding problem arising from overlapping downtimes across parts; gaming among suppliers and the consequences to realized performance; long-term, strategic product reliability investment vs. intermediate-term, tactical inventory decisions; investment in enhanced repair and logistics capabilities that would reduce lead times; alternative ownership and management scenarios; and many more. We are currently working on some of these issues. Finally, empirical verification of the insights gained from this paper will lead to more effective implementation of contract design and aid contract negotiations.

References

[1] Accenture. 2003. Service Management: Building Profits after the Sale. *Supply Chain Management Perspectives.*

[2] Bolton, P. and M. Dewatripont. 2005. *Contract Theory.* MIT Press.

[3] Cachon, G. 2003. Supply chain coordination with contracts. *Handbooks in Operations Research and Management Science: Supply Chain Management.* Edited by Steve Graves and Ton de Kok. North Holland.

[4] Chen, F. 2005. Salesforce Incentives, Market Information, and Production/Inventory Planning. *Management Science*, 51(1), 60-75.

[5] Chen, F. and A. Federgruen. 2000. Mean-Variance Analysis of Basic Inventory Models. Working paper, Columbia University.

[6] Cohen, M. A., V. Agrawal, and N. Agrawal. 2006. Achieving Breakthrough Service Delivery Through Dynamic Asset Deployment Strategies. *Interfaces.* Forthcoming.

[7] Cohen, M. A., P. Kamesam, P. Kleindorfer, H. Lee, A. Tekerian. 1990. OPTIMIZER: A Multi-Echelon Inventory System for Service Logistics Management. *Interfaces,* 20(1), 65-82.

[8] Corbett, C. J. 2001. Stochastic Inventory Systems in a Supply Chain with Asymmetric Information: Cycle Stocks, Safety Stocks, and Consignment Stock. *Operations Research*, 49(4), 487-500.

[9] Cummins, J. M. 1977. Incentive Contracting for National Defense: A Problem of Optimal Risk Sharing. *Bell Journal of Economics*, 8, 168-85.

[10] Department of Defense. 2003. Department of Defense Directive 5000.1. http://www.dtic.mil/whs/directives/corres/html/50001.htm.

[11] Defense Acquisition University. 2005. Defense Acquisition Guidebook. http://akss.dau.mil/dag.

[12] Defense Acquisition University. 2005. Performance Based Logistics: A Program Manager's Product Support Guide. http://www.dau.mil/pubs/misc/PBL_Guide.asp.

[13] Feeney, G. J. and C. C. Sherbrooke. 1966. The (s-1, s) Inventory Policy under Compound Poisson Demand. *Management Science*, 12(5), 391-411.

[14] Holmström, B. and P. Milgrom. 1991. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, and Organization*, 7, 24-52.

[15] Iyer, A., L. B. Schwarz and S. A. Zenios. 2005. A Principal-Agent Model for Product Specification and Production. *Management Science*, 51, 106-119.

[16] Kim, S.-H., M. A. Cohen, and S. Netessine. 2006. Technical Appendix to: Performance Contracting in After-Sales Service Supply Chains. Working paper, University of Pennsylvania. http://www.netessine.com.

[17] Laffont, J.-J. and J. Tirole. 1986. Using Cost Observation to Regulate Firms. *Journal of Political Economy*, 94(3), 614-641.

[18] Laffont, J.-J. and J. Tirole. 1993. *A Theory of Incentives in Procurement and Regulation*. MIT Press.

[19] Lutze, H. S. and O. Ozer. 2005. Promised Leadtime Contracts and Renegotiation Incentives Under Asymmetric Information. Working Paper, Stanford University.

[20] Macfarlan, W. G. and B. Mansir. 2004. Supporting the Warfighter through Performance-Based Contracting. *Defense Standardization Program Journal*, July/September. http://www.dsp.dla.mil/newsletters/std-news.html.

[21] Markowitz, H. 1952. Portfolio Selection. *Journal of Finance*, 7, 77-91.

[22] McAfee, R. P. and J. McMillan. 1986. Bidding for Contracts: A Principal-Agent Analysis. RAND Journal of Economics, 17(3), 326-338.

[23] Muckstadt, J. A. 1973. A Model for a Multi-Item, Multi-Echelon, Multi-Indenture Inventory System. *Management Science,* 20(4), 472-481.

[24] Muckstadt, J. A. 2005. *Analysis and Algorithms for Service Parts Supply Chains*. Springer.

[25] Plambeck, E. L. and T. A. Taylor. 2004. Partnership in a Dynamic Production System with Unobservable Actions and Noncontractible Output. Working paper, Columbia University.

[26] Plambeck, E. L. and S. A. Zenios. 2000. Performance-Based Incentives in a Dynamic Principal-Agent Model. *Manufacturing and Service Operations Management*, 2(3), 240-63.

[27] Scherer, F. M. 1964. The Theory of Contractual Incentives for Cost Reduction. *The Quarterly Journal of Economics*, 78(2), 257-80.

[28] Sherbrooke, C. C. 1968. METRIC: A Multi-Echelon Technique for Recoverable Item Control. *Operations Research*, 16, 122-41.

[29] Sherbrooke, C. C. 1971. An Evaluator for the Number of Operationally Ready Aircraft in a Multilevel Supply System. *Operations Research*, 19(3), 618-635.

[30] Sherbrooke, C. C. 1992. *Optimal Inventory Modeling of Systems: Multi-Echelon Techniques*. John Wiley and Sons.

[31] Slay, F. M. et al. 1996. Optimizing Spares Support: The Aircraft Sustainability Model. Logistics Management Institute, AF501MR1.

[32] So, K. C. and C. S. Tang. 2000. Modeling the Impact of an Outcome-Oriented Reimbursement Policy on Clinic, Patients, and Pharmaceutical Firms. *Management Science*, 46(7), 875-892.

[33] Su, X. and S. A. Zenios. 2005. Recipient Choice Can Address the Efficiency-Equity Trade-off in Kidney Transplantation: A Mechanism Design Model. Working paper, Stanford University.

[34] Van Mieghem, J. A. 2004. Risk-Averse Newsvendor Networks: Resource Sharing, Substitution, and Operational Hedging. Working paper, Northwestern University.

[35] Wang, Y., M. A. Cohen, and Y.-S. Zheng. 2000. A Two-Echelon Repairable Inventory System with Stocking-Center-Dependent Depot Replenishment Lead Times. *Management Science*, 46(11), 1441-1453.

Appendix

Proof of Proposition 1. We first prove that at the equilibrium, all (IR$_i$) constraints are binding, i.e., $w_i - (1-\alpha_i)(c_i s_i - a_i) - v_i E\left[B_i \,|\, s_i\right] - k_i a_i^2/2 - r_i(1-\alpha_i)^2 \text{Var}[\varepsilon_i]/2 - r_i v_i^2 \text{Var}[B_i \,|\, s_i]/2 = 0$ for all i. Suppose otherwise, i.e., that there exists j such that $w_j - (1-\alpha_j)(c_j s_j - a_j) - v_j E\left[B_j \,|\, s_j\right] - k_j a_j^2/2 - r_j(1-\alpha_j)^2 \text{Var}[\varepsilon_j]/2 - r_j v_j^2 \text{Var}[B_j \,|\, s_j]/2 > 0$. By reducing w_j by ϵ, the customer's utility (5) is increased by ϵ while the (AR) constraint is unaffected. This result allows us to transform $(\mathcal{A}_{FB})$ into

$$\begin{array}{cl} \min\limits_{\{\alpha_i, v_i, a_i, s_i\}} & \sum_{i=1}^n \left(c_i s_i - a_i + k_i a_i^2/2 + \left(r_0 \alpha_i^2 + r_i(1-\alpha_i)^2\right) \text{Var}[\varepsilon_i]/2 + (r_0 + r_i)\, v_i^2 \text{Var}[B_i \,|\, s_i]/2\right), \\ \text{s.t.} & \sum_{i=1}^n E\left[B_i \,|\, s_i\right] \le \widehat{B}_0 \,. \end{array}$$

Clearly, the objective function is minimized when $v_i = 0$ for all i. With this observation, the Lagrangian with the associated multiplier θ becomes

$$\begin{aligned} \mathcal{L} &= \sum_{i=1}^n \left(c_i s_i - a_i + k_i a_i^2/2 + \left(r_0 \alpha_i^2 + r_i(1-\alpha_i)^2\right) \text{Var}[\varepsilon_i]\right)/2 + \theta\left(\sum_{i=1}^n E\left[B_i \,|\, s_i\right] - \widehat{B}_0\right) \\ &= -\theta \widehat{B}_0 + \sum_{i=1}^n \left(c_i s_i + \theta E\left[B_i \,|\, s_i\right] - a_i + k_i a_i^2/2 + \left(r_0 \alpha_i^2 + r_i(1-\alpha_i)^2\right) \text{Var}[\varepsilon_i]/2\right). \end{aligned} \tag{14}$$

It is apparent that the minimization can be done separately for each supplier. Let $\mathcal{L}_i \equiv c_i s_i + \theta E\left[B_i \,|\, s_i\right] - a_i + k_i a_i^2/2 + \left(r_0 \alpha_i^2 + r_i(1-\alpha_i)^2\right) \text{Var}[\varepsilon_i]/2$. As the objective is a decreasing function of $\{s_i\}$ the optimal values are always at the corner and the (AR) constraint is binding, implying that $\theta > 0$. Note $\partial^2 \mathcal{L}_i / \partial a_i^2 = k_i > 0$, $\partial^2 \mathcal{L}_i / \partial s_i^2 = \theta f(s_i) > 0$, and $\partial^2 \mathcal{L}_i / \partial \alpha_i^2 = r_0 + r \geq 0$.In the absence of cross partial terms $\partial^2 \mathcal{L}_i / \partial a_i s_i = \partial^2 \mathcal{L}_i / \partial s_i \alpha_i = \partial^2 \mathcal{L}_i / \partial \alpha_i a_i = 0$, so the Hessian for supplier i is positive definite and hence the problem is convex, establishing the uniqueness of the equilibrium solution. (6), (7), and (9) are obtained from the first-order condition of supplier i. Clearly, the optimal s_i is a function of θ, which is determined from the (AR) constraint, as in (8). The supplier's profit and the customer's expenditures follow immediately. ■

Proof of Proposition 2. Note the following identities are needed:

$$
\begin{aligned}
d\text{Var}[B_i \,|\, s_i]/ds_i &= -2F_i(s_i)E[B_i \,|\, s_i] \leq 0, && (15)\\
d^2\text{Var}[B_i \,|\, s_i]/ds_i^2 &= -2f_i(s_i)E[B_i \,|\, s_i] + 2F_i(s_i)[1 - F_i(s_i)]. && (16)
\end{aligned}
$$

Let us drop the subscript i for notational convenience. Differentiating the supplier's expected utility function (4) with respect to s and substituting (15), we find that

$$
\partial U/\partial s = -(1-\alpha)c + v[1 - F(s)] + rv^2 F(s)E[B \,|\, s], \tag{17}
$$

which is greater than zero for all s if $\alpha = 1$ and $v > 0$, because U increases without bound in this case. If $\alpha < 1$ and $v[1 - F(0)] \geq (1-\alpha)c$, then $\partial U/\partial s \geq 0$ at $s = 0$, so U is nondecreasing initially. Notice also that $\lim_{s \to \infty} \partial U/\partial s = -(1-\alpha)c < 0$, so there exists at least one critical point on $[0, \infty)$. Setting $\partial U/\partial s = 0$, we get $E[B \,|\, s^*] = ((1-\alpha)c - v[1 - F(s^*)]) \,/\, \left(rv^2 F(s^*)\right)$. Substituting this result into the second derivative (use (16))

$$
\partial^2 U/\partial s^2 = -vf(s) + rv^2 f(s)E[B \,|\, s] - rv^2 F(s)[1 - F(s)] \tag{18}
$$

we obtain

$$
\partial^2 U/\partial s^2\big|_{s=s^*} = -[v - (1-\alpha)c]f(s^*)/F(s^*) - rv^2 F(s^*)[1 - F(s^*)] < 0,
$$

where the inequality follows from the condition $v \geq v[1 - F(0)] \geq (1-\alpha)c$. Since the second derivative is negative at every critical point, s^* cannot be a minimizer. Combining this result with $\partial U/\partial s|_{s=0} > 0$ and $\lim_{s \to \infty} \partial U/\partial s < 0$, we conclude that U has a unique maximizer. Optimal solutions follow from the first-order conditions. ■

Proof of Corollary 1. We drop the subscript i for notational convenience. After differentiating the first-order condition (17) implicitly with respect to r (optimal s is a function of r, i.e., $s^* = s(r)$) and collecting the terms we obtain

$$\frac{\partial s^*}{\partial r}=\frac{v^2F(s^*)E[B\,|\,s^*]}{vf(s^*)-rv^2f(s^*)E[B\,|\,s^*]+rv^2F(s^*)[1-F(s^*)]}.$$

Notice that the denominator has the sign opposite of that in (18). Hence $\partial s^*/\partial r>0$. Similarly,

$$\begin{aligned}\partial a^*/\partial\alpha &= -1/k<0, \qquad \partial a^*/\partial v=0,\\ \frac{\partial s^*}{\partial\alpha} &= \frac{c}{vf(s^*)-rv^2f(s^*)E[B\,|\,s^*]+rv^2F(s^*)[1-F(s^*)]}>0,\\ \frac{\partial s^*}{\partial v} &= \frac{[1-F(s^*)]+2rvF(s^*)E[B\,|\,s^*]}{vf(s^*)-rv^2f(s^*)E[B\,|\,s^*]+rv^2F(s^*)[1-F(s^*)]}>0.\end{aligned}$$

■

Proof of Proposition 3. With $r_0=r_1=...=r_n=0$, the customer's Lagrangian for supplier i becomes $\mathcal{L}_i(a_i,s_i,\theta)=c_is_i+\theta E\left[B_i\,|\,s_i\right]-a_i+k_ia_i^2/2$ and solutions are given by (6), (7), and (8). From the supplier's utility $U_i(a_i,s_i,w_i,v_i,\alpha_i)=w_i-(1-\alpha_i)(c_is_i-a_i)-v_iE\left[B_i\,|\,s_i\right]-k_ia_i^2/2$ it is clear that setting $\alpha_i=0$ and $v_i=\theta$ yields the same Lagrangian (with the reverse sign) as $\mathcal{L}_i$ plus a constant, reproducing the first-best solutions. The supplier's profit and the customer's expenditure follow immediately. ■

Proof of Lemma 1. Define $\gamma\equiv 4cF(s)E\left[B\,|\,s\right]/[1-F(s)]^2$ and note that (13) can be rewritten as

$$v(\alpha)=\frac{2c}{1-F(s)}\frac{1}{r\gamma}\left(-1+\sqrt{1+r\gamma(1-\alpha)}\right),$$

from which we obtain

$$v'(\alpha)=-\frac{c}{1-F(s)}\frac{1}{\sqrt{1+r\gamma(1-\alpha)}}, \qquad v''(\alpha)=-\frac{1}{2}\frac{c}{1-F(s)}\frac{r\gamma}{[1+r\gamma(1-\alpha)]^{3/2}}.$$

Differentiating the Lagrangian (12) and substituting $\partial(v^2)/\partial a=2v(a)v'(a)$ and $\partial^2(v^2)/\partial a^2=2(v'(a))^2+2v(a)v''(a)$ in it, we find that

$$\frac{\partial\mathcal{L}}{\partial\alpha}=\frac{\alpha}{k}+[(r_0+r)\alpha-r]\mathrm{Var}[\varepsilon]-\frac{2(r_0+r)c^2}{[1-F(s)]^2}\frac{1}{r\gamma}\left(1-\frac{1}{\sqrt{1+r\gamma(1-\alpha)}}\right)\mathrm{Var}[B\,|\,s],$$

$$\frac{\partial^2\mathcal{L}}{\partial\alpha^2}=\frac{1}{k}+(r_0+r)\mathrm{Var}[\varepsilon]+\frac{(r_0+r)c^2}{[1-F(s)]^2}\frac{1}{[1+r\gamma(1-\alpha)]^{3/2}}\mathrm{Var}[B\,|\,s]>0.$$

■

Proof of Proposition 5. We use the results in the proof of Lemma 1. Define

$$\begin{aligned}
\widetilde{\ell}_{FB}(\alpha) &\equiv [(r_0+r)\alpha - r]\mathrm{Var}[\varepsilon],\\
\widetilde{\ell}_{SO}(\alpha) &\equiv \frac{\alpha}{k} + [(r_0+r)\alpha - r]\mathrm{Var}[\varepsilon],\\
\widetilde{\ell}_{AO}(\alpha) &\equiv [(r_0+r)\alpha - r]\mathrm{Var}[\varepsilon] - \frac{2(r_0+r)c^2}{[1-F(s)]^2}\frac{1}{r\gamma}\left(1-(1+r\gamma(1-\alpha))^{-1/2}\right)\mathrm{Var}[B\,|\,s],\\
\widetilde{\ell}_{SB}(\alpha) &\equiv \frac{\partial\mathcal{L}}{\partial\alpha}(\alpha) = \frac{\alpha}{k} + [(r_0+r)\alpha - r]\mathrm{Var}[\varepsilon] - \frac{2(r_0+r)c^2}{[1-F(s)]^2}\frac{1}{r\gamma}\left(1-(1+r\gamma(1-\alpha))^{-1/2}\right)\mathrm{Var}[B\,|\,s].
\end{aligned}$$

α^{FB}, α^{SO}, α^{AO}, and α^{SB} are the solutions to $\widetilde{\ell}_{FB}(\alpha)=0$, $\widetilde{\ell}_{SO}(\alpha)=0$, $\widetilde{\ell}_{AO}(\alpha)=0$, and $\widetilde{\ell}_{SB}(\alpha)=0$, respectively. Observe $\widetilde{\ell}_{AO}(\alpha)\le\widetilde{\ell}_{SB}(\alpha)\le\widetilde{\ell}_{SO}(\alpha)$ for any α. Since $\widetilde{\ell}'_j(\alpha)>0$ for all j, $\alpha^{SO}\le\alpha^{SB}\le\alpha^{AO}$. In contrast, both $\widetilde{\ell}_{FB}(\alpha)\le\widetilde{\ell}_{SB}(\alpha)$ and $\widetilde{\ell}_{FB}(\alpha)\ge\widetilde{\ell}_{SB}(\alpha)$ are possible. To see this, substitute $\alpha^{FB}=r/(r_0+r)$ in $\widetilde{\ell}_{SB}$ to obtain

$$\widetilde{\ell}_{SB}(\alpha^{FB}) = \frac{1}{k}\frac{r}{r_0+r} - \frac{2(r_0+r)c^2}{[1-F(s)]^2}\frac{1}{r\gamma}\left[1-\left(1+\frac{r_0 r\gamma}{r_0+r}\right)^{-1/2}\right]\mathrm{Var}[B\,|\,s].$$

Let $\delta\equiv r/r_0$ and rewrite $\widetilde{\ell}_{SB}(\alpha^{FB})$ as a function of δ:

$$\ell_{SB}(\delta)\equiv\widetilde{\ell}_{SB}(\alpha^{FB}) = \frac{1}{k}\frac{\delta}{1+\delta} - \frac{2c^2}{[1-F(s)]^2}\frac{1}{r\gamma}\left(1+\frac{1}{\delta}\right)\left[1-\left(1+\frac{r\gamma}{1+\delta}\right)^{-1/2}\right]\mathrm{Var}[B\,|\,s].$$

Differentiating, we see that

$$\ell'_{SB}(\delta) = \frac{1}{k(1+\delta)^2} + \frac{2c^2\mathrm{Var}[B\,|\,s]}{[1-F(s)]^2 r\gamma}\left[\frac{1}{\delta^2}\left(1-\left(1+\frac{r\gamma}{1+\delta}\right)^{-1/2}\right) + \frac{(1+1/\delta)\,r\gamma}{2(1+\delta)^2}\left(1+\frac{r\gamma}{1+\delta}\right)^{-3/2}\right] > 0.$$

Hence, $\ell_{SB}(\delta)$ is increasing. Notice $\lim_{\delta\to0}\ell_{SB}(\delta)=-\infty$, $\lim_{\delta\to\infty}\ell_{SB}(\delta)=1/k$. Therefore there is a unique $\delta^\dagger$ such that $\ell_{SB}(\delta^\dagger)=0$. Since $\ell_{SB}(\delta)$ is increasing, $\ell_{SB}(\delta)=\widetilde{\ell}_{SB}(\alpha^{FB})>0$ for all $\delta>\delta^\dagger$, implying that $\alpha^{SB}<\alpha^{FB}$, since $\widetilde{\ell}_{SB}(\alpha)$ is also increasing. Likewise, $\alpha^{SB}>\alpha^{FB}$ for all $\delta<\delta^\dagger$.

Differentiating $\widetilde{\ell}_{SB}(\alpha^{SB})=0$ with respect to Var[ε] and collecting terms, we see that

$$\frac{d\alpha^{SB}}{d(Var[\varepsilon])} = \frac{r-(r_0+r)\alpha^{SB}}{\frac{1}{k}+(r_0+r)\mathrm{Var}[\varepsilon]+\frac{(r_0+r)c^2}{[1-F(s)]^2}\frac{1}{r\gamma}\frac{1}{[1+r\gamma(1-\alpha^{SB})]^{3/2}}\mathrm{Var}[B\,|\,s]}.$$

The numerator is negative if $\alpha^{SB}>\alpha^{FB}$, zero if $\alpha^{SB}=\alpha^{FB}$, and positive if $\alpha^{SB}<\alpha^{FB}$. The sign of $dv^{SB}/d(\mathrm{Var}[\varepsilon])$ is the reverse of that of $d\alpha^{SB}/d(\mathrm{Var}[\varepsilon])$ via (13). ∎

	$r_0 = 0.01$			$r_0 = 0.1$			$r_0 = 1$		
r_2	0.01	0.1	1	0.01	0.1	1	0.01	0.1	1
α_2^{FB}	0.500	0.909	0.990	0.091	0.500	0.909	0.010	0.091	0.500
α_1^{SB}	0.279	0.279	0.279	0.318	0.318	0.318	0.430	0.430	0.429
α_2^{SB}	0.060	0.279	0.755	0.184	0.318	0.726	0.436	0.430	0.539
v_1^{SB}	0.994	0.993	0.987	0.952	0.944	0.934	0.823	0.801	0.758
v_2^{SB}	1.408	0.993	0.288	1.215	0.944	0.317	0.826	0.801	0.490
a_1^{SB}	3.603	3.603	3.603	3.409	3.409	3.409	2.849	2.852	2.856
a_2^{SB}	4.699	3.603	1.224	4.082	3.409	1.371	2.820	2.852	2.303
s_1^{SB}	8.729	8.725	8.678	8.779	8.725	8.650	8.905	8.725	8.330
s_2^{SB}	8.722	8.725	8.773	8.672	8.725	8.802	8.559	8.725	9.151
$-U_0$	13.342	13.923	15.153	14.000	14.435	15.676	19.104	19.153	20.141

Table 4: Effects of changing r_2. Italics indicate symmetric parameters.

	$r_0 = 0.01$ ($\alpha_2^{FB} = 0.909$)			$r_0 = 0.1$ ($\alpha_2^{FB} = 0.5$)			$r_0 = 1$ ($\alpha_2^{FB} = 0.091$)		
$\mathrm{Var}[\varepsilon_2]$	1	10	100	1	10	100	1	10	100
α_1^{SB}	0.279	0.279	0.279	0.318	0.318	0.318	0.429	0.430	0.430
α_2^{SB}	0.172	0.279	0.650	0.268	0.318	0.442	0.654	0.430	0.156
v_1^{SB}	0.993	0.993	0.992	0.944	0.944	0.944	0.794	0.801	0.800
v_2^{SB}	1.127	0.993	0.507	1.007	0.944	0.785	0.505	0.801	1.149
a_1^{SB}	3.603	3.603	3.603	3.409	3.409	3.409	2.853	2.852	2.852
a_2^{SB}	4.138	3.603	1.752	3.658	3.409	2.790	1.731	2.852	4.221
s_1^{SB}	8.724	8.725	8.718	8.725	8.725	8.723	8.670	8.725	8.715
s_2^{SB}	8.727	8.725	8.733	8.726	8.725	8.728	8.782	8.725	8.735
$-U_0$	13.653	13.923	15.139	14.172	14.435	16.780	17.794	19.153	24.341

Table 5: Effects of changing $Var[\varepsilon_2]$ when $r_1 = r_2 = 0.1$. Italics indicate symmetric parameters.

$\widehat{B}_0$	1	2	3	4	5	6	7
α_1^{SB}	0.421	0.404	0.398	0.396	0.396	0.397	0.399
α_2^{SB}	0.592	0.604	0.614	0.624	0.634	0.644	0.653
v_1^{SB}	1.678	1.147	0.935	0.820	0.748	0.699	0.666
v_2^{SB}	0.792	0.557	0.463	0.412	0.379	0.357	0.341
a_1^{SB}	2.896	2.980	3.011	3.021	3.021	3.015	3.006
a_2^{SB}	2.039	1.981	1.929	1.879	1.830	1.782	1.737
s_1^{SB}	11.852	10.384	9.355	8.511	7.765	7.078	6.425
s_2^{SB}	12.205	10.752	9.751	8.948	8.259	7.644	7.083
$(s_2^{SB} - s_1^{SB})/s_1^{FB}$	2.94%	3.48%	4.15%	5.00%	6.17%	7.69%	9.75%
IIR_1	11.852	10.384	9.355	8.511	7.765	7.078	6.425
IIR_2	12.205	10.752	9.751	8.948	8.259	7.644	7.083
NCR_1	2.057	2.092	2.104	2.108	2.108	2.106	2.102
NCR_2	1.623	1.589	1.557	1.526	1.495	1.465	1.435
CRP_1	0.610	0.586	0.577	0.574	0.574	0.576	0.578
CRP_2	1.708	1.696	1.687	1.680	1.675	1.671	1.668
PRP_1	1.252	1.160	1.116	1.090	1.074	1.062	1.055
PRP_2	0.574	0.587	0.600	0.611	0.621	0.628	0.632
$-U_0$	24.521	21.484	19.425	17.781	16.364	15.088	13.905

Table 6: Effects of changing $\hat{B}_0$, with $r_0 = 0.5$, $r_1 = 0.1$ and $r_2 = 1$. IIR stands for investment in resources and is equal to $c_i s_i^{SB}$. NCR is $-a_i^{SB} + \frac{1}{2}k_i(a_i^{SB})^2$, the net cost reduction. CRP is the residual cost risk premium, $\frac{1}{2}(r_0(\alpha_i^{SB})^2 + r_i(1-\alpha_i^{SB})^2)Var[\varepsilon_i]$, and PRP is the residual performance risk premium, $\frac{1}{2}(r_0 + r_i)(v_i^{SB})^2 Var[B_i|s_i^{SB}]$.