IT Implementation Contract Design: Analytical and Experimental Investigation of IT Value, Learning, and Contract Structure

D. J. Wu
Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia 30332,
dj.wu@scheller.gatech.edu

Min Ding
Smeal College of Business, Pennsylvania State University, University Park, Pennsylvania 16802; and
School of Management, Fudan University, Shanghai, P. R. China, 200433, minding@psu.edu

Lorin M. Hitt
The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104,
lhitt@wharton.upenn.edu

This article analytically and experimentally investigates how firms can best capture the business value of information technology (IT) investments through IT contract design. Using a small sample of outsourcing contracts for enterprise information technology (EIT) projects in several industries—coupled with reviews of contracts used by a major enterprise software maker—the authors determine the common provisions and structural characteristics of EIT contracts. The authors use these characteristics to develop an analytical model of optimal contract design with principal–agent techniques. The model captures a set of key characteristics of EIT contracts, including a staged, multiperiod project structure; learning; probabilistic binary outcomes; variable fee structures; possibly risk-averse agents; and implementation risks. The model characterizes conditions under which multistage contracts enable clients to create and capture greater project value than single-stage projects, and how project staging enables firms to reduce project risks, capture learning benefits, and increase development effort. Finally, the authors use controlled laboratory experiments to complement their analytical approaches and demonstrate robustness of their key findings.

1. Motivation

Large-scale enterprise IT (EIT) implementation projects (e.g., Chellappa and Saraf 2010, Hitt et al. 2002) are risky by nature (Haines and Goodhue 2003, McFarlan 1981, O'Leary 2002). According to a report by a leading enterprise resource planning (ERP) vendor, 51% of projects suffered from unforeseen implementation issues, 53% of projects exceeded cost estimates, 83% of projects were delivered late or over budget, 42% of projects had incomplete features or functions, and 40% of projects failed to achieve their business cases (Pike 2006). Empirical evidence in an ERP context suggests that managers and investors perceive particularly high risks for EIT projects (e.g., Hitt et al. 2002). EIT projects typically get outsourced to consulting companies for implementation and maintenance. Such projects are notoriously complicated and place burdens on the design of IT outsourcing relationships, especially on the contracts that codify these relationships. Nearly all large firms have, or are in the process of implementing, large-scale enterprise systems, and it is not unusual for these projects to include various contracting opportunities spanning multiple years (e.g., O'Leary 2002). Contracts provide the primary means of IT governance (Nolan and McFarlan 2005), codifying client–vendor relationships in IT outsourcing agreements (Clemons et al. 2001). Well-designed contracts can help manage the problems of ex-ante incomplete information (e.g., requirements, client characteristics, or vendor capabilities) and provide a framework for measuring performance, providing incentives, and managing technical, business, and managerial risks. Yet considerable evidence indicates that many outsourcing agreements prove difficult to manage, as reflected in the serious problems observed in some major ERP implementations (e.g., McAfee 2003, Mendelson 2000, Scott 1999).

EIT implementation contracts tend to be similar across different projects because they rely on the commoditization and standardization of business processes (Davenport 2005). In addition, the industry consists of few dominant software vendors and implementation consultants who generally share common business practices. Projects also contain features common to regular IT contracts, though on a larger scale; such as the extensive use of outside consultants, the use of packaged software, and largely observable initial results (i.e., the client can either "go live" with the system or not).

The similarity of ERP contracts across installations and vendors makes it possible to capture the variation of ERP contracts along relatively few dimensions. In this paper, we focus on one particular characteristic that creates considerable variation across projects—the choice between structuring the project as a single-stage "big bang" implementation without any intermediate decision points, and a multistage incremental "rollout" with scheduled intermediate deliverables enabling a "continue or terminate" decision. Thus, a critical question in structuring an ERP project is how long to make each stage, with the possibility that it may be optimal to make the first stage the entire project.

The nature of ERP projects clearly makes stage length a strategic choice. Implementations can be done in the form of a pilot implementation, which is then subsequently rolled out geographically in an identical form to multiple sites. Alternatively, ERP projects can be implemented by module—with core functions implemented over multiple sites (e.g., manufacturing planning, procurement, human resources), followed by additional modules that provide supplemental functionality (e.g., supply chain management, customer relationship management). A multistage contract enables the effective use of information gathered during the project, such as midterm project evaluations, that may reduce subsequent period project risks, improve vendor incentives,[1] or enable early terminations of unsuccessful projects. This phased approach also offers a means to assess best practices and disseminate them to the project team, so teams can improve their performance through learning. Because process redesign and testing, training, and deployment can entail an estimated 12%–15% of total project expenditures in a typical EIT project (Brynjolfsson et al. 2006), client-specific learning can have a significant impact on project value (e.g., Hitt et al. 2002, McAfee 2003). ERP vendors understand

the benefits of staging and having a flexible project scope:

> Implementation of SAP software is a process that often involves a significant resource commitment by our customers and it is subject to a number of significant risks over which we have little or no control.... Our customers now increasingly follow modular project approaches to optimize their IT environment. They embark on sequentially integrated individual projects with a comparatively low-risk profile to realize specific potential improvement instead of pursuing highly complex resource-intensive "big bang" projects to implement an all-embracing IT landscape.
> (SAP 2008 annual report, p. 123)

Subdividing a project is not without costs, which may explain the use of "big bang" projects in practice. Subdividing a project may increase coordination costs between stages and delay the implementation of later stages, thus reducing or delaying realization of value (e.g., some mission-critical enterprise systems may have more value by going live simultaneously rather than sequentially). Staged implementation may also lead to overinvestments during early periods, as advancing to the next period requires success in the previous period.

Our model focuses specifically on the choice of project staging. We consider a project of fixed duration that can be divided into two stages and outsourced to an implementation vendor. Allowing for different learning processes and project risks, we derive the optimal contract which specifies vendor compensation (fixed and outcome-based payments for each stage) and the optimal stage length. All else being equal, combining the optimal payment structure with a multistage contract allows the client to capture greater overall IT project value, because it mitigates project risks and increases learning benefits by prompting the implementation team to exert greater efforts in early stages. However, under some conditions, such as when there is limited learning, "big bang" full projects can be optimal. These findings are robust to several model extensions. We use controlled laboratory experiments involving students and IT executives to complement our analytical approaches and demonstrate further robustness of our key findings.

All proofs appear in the online supplement of this paper (available at http://dx.doi.org/10.1287/isre.1120.0448). Due to space limitations, we have omitted details of the laboratory experiments but they are available from the authors upon request.

2. Literature Review
Our work draws from three streams of research. We briefly review each, linking key elements of our model to the literature.

---

[1] Haines and Goodhue (2003) document three cases of moral hazard in ERP implementations, including one case of an outcome failure (production showdown for almost a month) that they attribute to consultant moral hazard behavior.

2.1. Learning and Dynamic Production Function

The first stream of related literature involves the dynamic nature of production processes as firms increase their productive capabilities through experience. Levy (1965) was among the first to model these processes explicitly, using an exponential production function model that enables a firm to accumulate experience in its workforce through training and production activity until it achieves target productivity (or the full potential of its production technology). Gaimon (1997) considers the underlying processes that may characterize the productivity of IT-enabled knowledge work and describes the desirable attributes of a production function in this setting. Gaimon's analysis suggests that an exponential functional form that incorporates a learning process meets all the required criteria and appears superior to some standard alternatives for modeling IT implementation projects. Case-based research on ERP systems is consistent with a time-dependent learning process, both before and after implementation (McAfee 2002).

We adopt this dynamic production process (learning) view of IT implementation and, specifically, the exponential functional form that characterizes the process (Levy 1965). This structure incorporates decreasing returns, which dates back to Brooks (1975) in the field of software development, and has been well documented in project management literature as well (e.g., Loch and Kavadias 2002), especially in the IT project management literature (e.g., Banker et al. 1998, Barry et al. 2002, Kirsch 2000, McFarlan 1981). We extend Levy's basic model to capture the key factors in the IT implementation context, including the role of skill, effort, and project size in determining project outcomes, in the presence of both controllable and uncontrollable project risks. Researchers also note that training and learning represent important components of project risk management (Anderson 2001, Banker et al. 1998, Barry et al. 2002, Umble et al. 2003), both before and during a project (Anderson 2001, Gaimon et al. 2011). Thus, we extend and reinterpret Levy's workforce skill–outcome model by allowing the relationship between effort and project success rate to vary by vendor capability, which can evolve through learning.

2.2. Multi-Period Moral Hazard

The second stream of related research entails the well-established literature on principal–agent formalisms, which considers the general problem of providing incentives in a variety of settings (e.g., Holmstrom 1982; McAfee and McMillan 1986, 1987). Typically, to mitigate the moral hazard problem that results from unobservable effort by the agent, principals implement incentive contracts that compensate agents on the basis of observable outcomes. The literature has well characterized the properties of the single-period principal–agent problem. However, we know significantly less about multiperiod (or finite-horizon) moral hazard games (e.g., Chiappori et al. 1994), the setting that seems to naturally characterize large-scale IT implementations. Prior works by Lambert (1983) and Rogerson (1985) suggest a special role of memory in finite horizon moral hazard models, in contrast with infinite horizon models where compensation can be based on long-run summary statistics (e.g., Holmstrom and Milgrom 1987; Radner 1981, 1985). See Chiappori et al. (1994) for a survey of this literature.

We apply the most recent developments in controlling dynamic moral hazard in contract theory in the EIT contract setting with small variations, in order to capture the details and key project characteristics observed from actual EIT contracts, such as the linear fee structure. Essential to EIT projects are linkages between periods, which have rarely been considered in the existing literature. In particular, we consider the potential for change in the effort–output relationship through agent learning over time and risk reduction through information updating over the course of the project.

2.3. IT Outsourcing and Contracting

The final stream of related literature pertains to outsourcing risk management (e.g., Clemons et al. 2001) and software development contracting (e.g., Richmond and Seidmann 1993, Richmond et al. 1992, Wang et al. 1997, Whang 1992), as well as the broader literature on IT outsourcing, which has taken a much more qualitative evaluation approach (e.g., DiRomualdo and Gurbaxani 1998, Lacity and Hirschheim 1993, Lacity and Willcocks 1998).

Our analysis complements research that emphasizes monitoring (Nolan and McFarlan 2005) and incentives (Choudhury and Sabherwal 2003) in IT contracting, as well as the uncertainty reduction achieved through information updating over the course of a project (Snir and Hitt 2004, Whang 1992).

Through analyses of actual ERP contracts, we identify at least two major structural dimensions that capture much of the variation in project structure.[2] First, though firms purchase ERP software externally, they must decide whether to perform the implementation in-house (insource) or contract with a service provider (outsource). Second, an outsourced contract can be a single-stage "big bang" implementation or a series

---

[2] Over a period of several years (2000–2006), students of two major U.S. universities searched and analyzed publicly available ERP contracts for course credits, drawing on sources provided by clients, ERP consultants, and software vendors. In addition, the research team conducted site visits to examine real-world EIT projects and a set of real-life contracts from a major ERP vendor.

of sequentially interlinked subprojects that give the client a termination decision at the end of each period.

Beyond these major decisions, the contracts principally differ in their use of fixed-fee versus incentive payments and the extent to which the vendor makes relationship-specific investments (e.g., training, facilities). We observed that the variable portions of payments may be structured in several ways. Some contracts include specific bonus/penalty clauses, others include negotiated rate increases or discounts for future work, and some guarantee a preferential position ("right of first refusal") to the vendor in future bidding. In one set of contracts, we observed links of productivity and performance with ERP implementation in the oil and gas industry, in which pre-ERP performance serves as a benchmark for the revenue-sharing agreement associated with post-ERP implementation. These structures can be captured using a fixed-fee plus performance-based variable incentive structure—to the extent that their objectives are to provide incentives to the vendor. These observations are consistent with incentives observed in practice pertaining to IT service consulting and development (e.g., James 1998) or in research pertaining to after-sales service support (Cohen et al. 2006).

In this paper, we focus on the firm's optimal contract choice regarding the two key structural dimensions—payment structure and sizing—in the presence of agent learning. As a benchmark, we also compute the optimal sizing and value of an insourced project (where efforts are observable and there are no moral hazard issues), and compare and contrast the insourced case with the outsourced case.

Together, these arguments naturally suggest considering enterprise software contracting in a multiperiod setting, such that the outcome (success or failure) in the first period can influence continuation to a second period (e.g., Pike 2006). We formally present a model that integrates various isolated elements of the three streams of research to capture the aforementioned salient features of EIT implementation. Finally, we study the properties of the EIT project value function, and test model assumptions and key predictions experimentally.

3. Model

We assume that a risk-neutral principal contracts with a risk-neutral or a risk-averse agent for an enterprise IT project that generates a revenue of Q. The principal can divide the project into two sequential periods, with the outcome of each period randomly determined but observable at the end of each period. The outcome is assumed to take binary values of 1 (success) or 0 (failure). The contract specifies that continuation of the project to the second period project is conditional on success in the first period. The agent

is assumed to be committed to complete the project if desired by the principal, consistent with observed contracting practice. The key notation for our model appears as follows.

Q Total revenue created by the EIT project

s_i Size of the EIT project in period i ($s_1 = \alpha$, $s_2 = 1 - \alpha$)

$R_i \equiv s_i Q$ Revenue of period i ($i = 1, 2$)

$a_i + b_i R_i$ Two-part wage where a_i is the fixed fee, $b_i R_i$ is revenue sharing ($i = 1, 2$)

$u(x)$ Agent's utility of profit

γ Risk tolerance of the agent

r Interest rate

λ Discrete case second period fixed discount factor. In the continuous case, $\lambda_1 = e^{-rs_1}$, $\lambda = \lambda_2 = e^{-rs_2}$, where r is the interest rate

$\bar{P}_i$ General IT implementation capability in period i ($i = 1, 2$)

$(1 - \bar{P}_i)$ Inherent IT project risk in period i ($i = 1, 2$)

p_i Production function in period i ($p_i = \bar{P}_i(1 - e^{-\beta_i(x_i/s_i)})$, $i = 1, 2$)

β_i Average skill (or expertise) in period i ($i = 1, 2$)

x_i Effort in period i ($x_i = -(s_i/\beta_i) \cdot \ln(1 - p_i/\bar{P}_i)$, $i = 1, 2$)

Δ_i $\Delta_i \equiv \bar{P}_i Q - 1/\beta_i - \ln(\beta_i \bar{P}_i Q)/\beta_i$ ($i = 1, 2$)

π_i^0 Period i reservation profit (normalized to 0)

π^0 Reservation profit over the two-period horizon (normalized to 0)

Π_i Principal's expected profit from period i

V Total EIT project value captured by the principal.

We further assume period i ($i = 1, 2$) subproject generates revenue of $R_i = s_i Q$ if successful, where s_i is the size (scope, duration or time) of the subproject in that period, otherwise $R_i = 0$. We normalize the overall project size to 1 (i.e., $s_1 + s_2 \equiv 1$) so we need only one variable $s_1 \equiv \alpha$ to capture sizing of the full project (because $s_2 = 1 - \alpha$). The case when $\alpha = 1$ illustrates that the principal contracts the full project to the agent in a single stage. We denote $\underline{\alpha}$ the smallest feasible project size (e.g., the core module) so that $0 < \underline{\alpha} \le \alpha \le 1$. In practice, α represents the principal's choice of project staging, including a full project or any subset of sequentially interlinked incremental phases, such as planning and design, construction and implementation, maintenance, support, and ongoing services related to the project (e.g., Markus et al. 2000, McAfee 2003). We assume stage 1 must include the core development, but concurrently other components may also be developed.[3] In a software-as-a-service

---

[3] We thank a referee for this clarification, which makes the scope of stage 1 endogenous.

context (Chou 2010), stage length represents the time between contract renewals. For simplicity, we assume a continuous α.

The principal's pure strategy is to specify a wage function in a long-term contract $[a_i, b_i; s_i]$, $i = 1, 2$, which takes the linear form of $a_i + b_i R_i$. Here a_i is a fixed fee, $b_i > 0$ is a revenue sharing factor contingent on outcome success, otherwise $b_i = 0$. The agent's pure strategy is the mapping from the wage function to the agent's action or effort x_i. For tractability as well as consistency with observed contracting practice (see §2.3), we restrict our attention to two-part linear contracts rather than abstract reward functions in the extant literature (e.g., Lambert 1983, Rogerson 1985). Our two-part linear contracts, however, are quite general. For instance, parameter choices can yield a "big bang" ($s_1 = \alpha = 1$) or multistage ($s_1 < 1$) project structure, fixed fee ($b_i = 0$, $i = 1, 2$) contracts or incentive ($b_i > 0$, $i = 1, 2$) contracts, with or without upfront vendor investments ($a_i \geq 0$, $i = 1, 2$), contracts in which a vendor posts bonds or makes fixed co-investments in project preparations ($a_i \leq 0$, $i = 1, 2$), and various combinations of these characteristics.

We extend existing moral hazard models by incorporating learning. The existence of effort-dependent learning is perhaps the most important, and the primary assumption that distinguishes our analysis from other principal-agent analyses. Following Levy (1965) and Gaimon (1997), we initially adopt an exponential learning function that relates outcome in each stage (p_i) with inherent project risk ($1 - \bar{P}_i$), skill β_i and project stage size s_i.

ASSUMPTION A0 (DYNAMIC PRODUCTION FUNCTION). $p_1(x_1, s_1, \bar{P}_1, \beta_1) = \bar{P}_1(1 - e^{-\beta_1(x_1/s_1)})$, and $p_2(x_2, s_2, \bar{P}_2, \beta_2 \mid period\ 1\ succeeds) = \bar{P}_2(1 - e^{-\beta_2(x_2/s_2)})$.

Essentially, our dynamic production function captures two types of risks: inherent project risks ($1 - \bar{P}_i$), which represent technical, business, or managerial challenges that are part of the project (e.g., O'Leary 2002), and controllable risks $e^{-\beta_i(x_i/s_i)}$, which are influenced by agent effort x_i, expertise β_i, and project size s_i. For a given project size at each period, because inherent project risks and controllable risks are statistically independent, outcome success of the EIT project depends on the successful removal of both risks, leading to the multiplicative form of p_i we are using.

The information structure of the game is as follows. We assume the agent's expertise levels β_i are common knowledge. Because both project value and agent expertise are known to the client, the potential adverse selection issues (Snir and Hitt 2004, Whang 1992) have presumably been resolved, and our model focuses purely on moral hazard
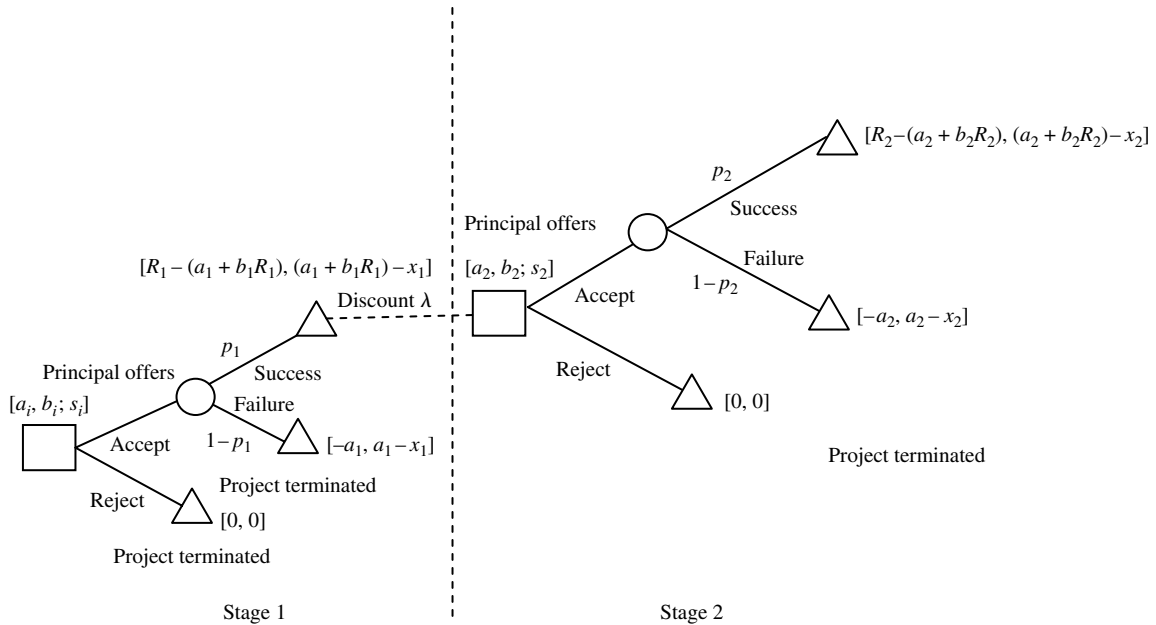
(e.g., Holmstrom and Milgrom 1987).[4, 5] At the end of each period, the binary outcomes are observable and contractible; in essence, the outcome is whether the principal accepts or rejects the agent's delivery. However, the agent's efforts are not observable and cannot be "reverse engineered" from observable data because of the random component of the outcome. Conditional on a successful first-period, some inherent project risks can be reduced such that $1 - \bar{P}_2 \leq 1 - \bar{P}_1$, perhaps through business process redesign. Consequently, both the principal and the vendor share a common updated belief such that $\bar{P}_2 \geq \bar{P}_1$. We discuss the case when updating is affected by project size α in a model extension, but we initially treat these capability factors as constants. The possibility of learning between the two periods derives from learning-by-doing or deliberate investments in team skill and project-specific risk management capabilities (e.g., post-project audit, staff training, or staff reallocation). Learning changes the relationship between effort and project outcomes, and the updated skill level is common knowledge.

The nature of the contracts and the structure of the industry make it unlikely that the contract will be subject to renegotiation after the first stage. We assume that the agent is bound to complete the project if desired by the principal (a practice we have observed) and derive the optimal contract under agent individual rationality in both periods. Provided that project learning is project specific, the vendor's outside options do not change as the project progresses and thus the vendor cannot credibly threaten to terminate a successful contract even if it was contractually permissible. In optimum, the principal may desire to continue a failed project as the principal knows this was due to uncontrollable risks and not agent under performance, but is unable to do so because a failed first stage in most IT implementations prevents progressing to later stages. For instance, if the

---

[4] Because a limited number of vendors engage in large-scale EIT work, we presume the client can observe the vendor's history to infer its skill ex-ante, and monitoring and coproduction during the project enable assessments of expertise changes during the project. When initial expertise is unknown, the client can engage in a selection process or a separate pilot project to determine vendor skill (e.g., Snir and Hitt 2004).

[5] It is worthwhile to compare and contrast the real options approach in risk management with the model we are developing, which focuses on staging. Staging has the same structure as a real options problem, which is a deferred decision. In our context, the uncertainty is solely related to the realization of contract risks, and the contract is not terminated because something is learned about the setting or the vendor but because an intermediate failure prevents continuation. It is interesting to note that staging does limit the principal's risk in the presence of adverse selection, but that is not an explicit component of our model. In our contracts, both parties understand and agree ex-ante that there will be no second period project if the first period fails.

Figure 1   Decision Tree

core module of an ERP installation is not operational, supplemental models cannot be installed. This is a different structure than many prior repeated agency models which has the agent performing similar tasks in each stage and where principal commitment is often enforced with auxiliary assumptions such as reputational concerns.[6]

We employ a sub-game perfect equilibrium (SPE) solution concept (e.g., Bolton and Dewatripont 2005, p. 421; Chiappori et al. 1994; Lambert 1983). Figure 1 depicts the decision tree of the principal and the agent.

4.   Impact of Contract Structure

We assume the agent is risk-neutral. Later on we extend the model to the case when the agent is risk-averse. When both the principal and the agent are risk-neutral, the linear contract form has a useful property that it is optimal among all possible contract forms (e.g., Holmstrom and Milgrom 1987).

We first consider the case (see §4.1) when the principal wishes to structure the EIT project as a single-stage "big bang" contract. We then consider the two-stage case (see §4.2).

4.1.   Optimal Single-Stage Contract

For simplicity, we first assume no discounting for the single-stage case such that $\lambda = 1$. Later we extend to the case when there is a discount. The

principal's problem is:

$$\max_{a_1, b_1} \big\{ V(a_1, b_1; s_1 = 1)$$
$$= \Pi_1 = [p_1(R_1 - a_1 - b_1 R_1) + (1 - p_1)(-a_1)] \big\},$$

subject to

(i) $\pi(x_1^*) = p_1(a_1 + b_1 R_1 - x_1^*) + (1 - p_1)(a_1 - x_1^*) \geq \pi^0$,
    (Agent IR)

(ii) x_1^* solves $\max_{x_1} \{ p_1(a_1 + b_1 R_1 - x_1) + (1 - p_1) \cdot (a_1 - x_1) \}$.
    (Agent IC)

The first constraint (i) requires that the agent's expected payoff must exceed the agent's reservation profit π^0 to satisfy the agent's individual rationality requirement. The second constraint (ii) satisfies the agent's incentive compatibility requirement.

Solving the principal's problem, we obtain the following Lemma 1. For convenience of exposition, we define $\bar{P}_i Q - 1/\beta_i - \ln(\beta_i \bar{P}_i Q)/\beta_i \equiv \Delta_i$.

LEMMA 1. *Assume agent is risk-neutral. Optimal single-stage contract (a_1^*, b_1^*), agent effort (x_1^*) and optimal IT project value (V^*) are given as follows:*

$$a_1^* = -\left[\bar{P}_1 Q - \frac{1}{\beta_1} - \frac{\ln(\beta_1 \bar{P}_1 Q)}{\beta_1} \right] + \pi^0 = -\Delta_1 + \pi^0,$$

$$b_1^* = 1,$$

$$x_1^* = \frac{\ln(\beta_1 \bar{P}_1 Q)}{\beta_1},$$

$$V^* = -a_1^* = \left[\bar{P}_1 Q - \frac{1}{\beta_1} - \frac{\ln(\beta_1 \bar{P}_1 Q)}{\beta_1} \right] - \pi^0 = \Delta_1 - \pi^0.$$

It is straightforward to verify that the above optimal solution is efficient, as it is identical to the solution when the agent effort is observable (therefore removing moral hazard) or if the principal can do the project alone (assuming they have the expertise).

4.2. Optimal Two-Stage Contract

We now consider the principal's two-stage contracting problem. We solve the model in two steps. First, given any project sizing α, with $0 < \underline{\alpha} \leq \alpha = s_1 \leq 1$ (thus $s_2 = 1 - \alpha$), we solve for the optimal fee structures (a_i^*, b_i^*) at each period i taking into account how different incentive payments affect agent effort and participation. Second, we solve for the optimal sizing $\alpha^* = s_1^*$ (with $s_2^* = 1 - \alpha^*$).

To obtain optimal fee structures (a_i^*, b_i^*), we use backward induction (i.e., dynamic programming). At the beginning of period 2, the principal updates the contract (a_2, b_2) after observing the agent's first-period performance:

$$\max_{a_2, b_2} \Pi_2 = p_2[R_2 - (a_2 + b_2 R_2)] + (1 - p_2)(-a_2),$$

subject to

(i) $\pi_2(x_2^*) = p_2(a_2 + b_2 R_2 - x_2^*) + (1 - p_2)(a_2 - x_2^*) \geq \pi_2^0$, (Agent IR)

(ii) x_2^* solves $\max_{x_2}\{\pi_2(x_2) = [p_2(a_2 + b_2 R_2 - x_2) + (1 - p_2)(a_2 - x_2)]\}$. (Agent IC)

The objective function is the principal's expected profit in the second period, given the updated information of $\bar{P}_2 \geq \bar{P}_1$ and the agent's updated expertise $\beta_2 \geq \beta_1$, so that

$$p_2(x_2, s_2, \bar{P}_2, \beta_2 \mid \text{period 1 succeeds}) = \bar{P}_2(1 - e^{-\beta_2(x_2/s_2)}).$$

Solving, the principal's optimal second period strategy (a_2^*, b_2^*), and the agent's optimal effort (x_2^*) are:

$$a_2^* = -s_2\left[\bar{P}_2 Q - \frac{1}{\beta_2} - \frac{\ln(\beta_2 \bar{P}_2 Q)}{\beta_2}\right] + \pi_2^0, \quad (1)$$

$$b_2^* = 1, \quad (2)$$

$$x_2^* = s_2 \frac{\ln(\beta_2 \bar{P}_2 Q)}{\beta_2}. \quad (3)$$

Similarly, the principal's problem at time 0 (i.e., the beginning of period 1) is:

$$\max_{a_1, b_1} V(a_1, b_1; s_1, s_2, a_2^*, b_2^*)$$

$$= \Pi_1 + p_1 \lambda \Pi_2^*$$

$$= [p_1(R_1 - a_1 - b_1 R_1) + (1 - p_1)(-a_1)] + p_1 \lambda \Pi_2^*,$$

subject to

(i) $\pi(x_1^*) = p_1(a_1 + b_1 R_1 - x_1^*) + (1 - p_1)(a_1 - x_1^*) + p_1 \lambda \pi_2^* \geq \pi^0$, (Agent IR)

(ii) x_1^* solves $\max_{x_1}\{p_1(a_1 + b_1 R_1 - x_1) + (1 - p_1) \cdot (a_1 - x_1) + p_1 \lambda \pi_2^*\}$. (Agent IC)

Solving, the principal's first period optimal strategy (a_1^*, b_1^*), and the agent's optimal effort in response (x_1^*) are:

$$a_1^* = -(s_1 \Delta_1 + \bar{P}_1 \lambda s_2 \Delta_2) + \frac{s_1}{\beta_1}\ln\left(1 + \frac{\lambda s_2 \Delta_2}{s_1 Q}\right) + \pi^0, \quad (4)$$

$$b_1^* = 1 + \frac{\lambda(s_2 \Delta_2 - \pi_2^0)}{s_1 Q}, \quad (5)$$

$$x_1^* = \frac{s_1}{\beta_1}\ln\frac{\beta_1 \bar{P}_1(s_1 Q + \lambda s_2 \Delta_2)}{s_1}. \quad (6)$$

Consequently, the principal's total IT project value (or IT value function) is

$$V^* = \Pi_1^* + p_1 \lambda \Pi_2^* = -a_1^*$$

$$= (s_1 \Delta_1 + \bar{P}_1 \lambda s_2 \Delta_2) - \frac{s_1}{\beta_1}\ln\left(1 + \frac{\lambda s_2 \Delta_2}{s_1 Q}\right) - \pi^0. \quad (7)$$

The IT value function as given in Equation (7) highlights the central trade-off in sizing (i.e., determining optimal s_i^*) in a two-stage project—building higher expertise ($\beta_2 \geq \beta_1$) and increasing implementation capability ($\bar{P}_2 \geq \bar{P}_1$) for the second stage (via learning in the first stage), versus investing additional effort (third term in Equation (7)) to increase the chance of continuation to the second stage of the project. There is also a trade-off in timing; while a larger first-stage project size increases second-stage expertise and capability, it shrinks time left ($s_2^* = 1 - s_1^*$) to harvest such benefits ($\Delta_2 \geq \Delta_1$, or Δ_i is increasing in both β_i and $\bar{P}_i$). We shall formally characterize the properties of the IT value function later via Theorem 1 and its proof.

The optimal two-stage contract as defined in Equations (1), (2), (4), (5) is equivalent to the principal selling the project to the agent at each period. The agent nets exactly its reservation profit because both IR and IC constraints are binding (this property is later exploited in our lab experiments). Note that Δ_i ($i = 1, 2$) can be interpreted as a kind of "unit price."[7] Specifically, Δ_2 is a second-stage "unit price" since the agent pays $s_2 \Delta_2 - \pi_2^0$. Similarly, Δ_1 can be interpreted as a first-stage "unit price."

Further, we show in Proposition 1 that the above optimal contract is *viable* (Whang 1992). A contract is said to be viable if it satisfies the following four properties (Whang 1992): (1) Efficiency: It induces the same equilibrium decisions as the optimal in-house implementation at each period; (2) Pareto Optimality: The combined equilibrium payoff to the contracting parties is the same as in in-house implementation; (3) Incentive Compatibility: The principal reports the true value of the system, which induces the agent to exert an effort that maximizes its own profit; and (4) Ex-ante Incentive Rationality: Both contracting

---

[7] We thank an anonymous referee for this insight.

parties have non-negative expected payoffs at the time of contracting, so they voluntarily sign the contract.

PROPOSITION 1. *Assume agent is risk-neutral. Given any sizing of an IT project* s_i ($i = 1, 2$),

(i) *The linear contract form is optimal among all possible contract forms.*

(ii) *There exists a unique SPE for the principal's problem.*

(iii) *The optimal contract* $[a_i^*, b_i^*; s_i]$ ($i = 1, 2$), *as specified in Equations* (1), (2), (4), (5), *is viable.*

Proposition 1 contributes to the literature by establishing the structure of the optimal contract capturing unique characteristics of EIT projects, and by extending multiperiod agency models to a setting where there is agent learning across project periods. Proposition 2, as follows, characterizes the comparative statics of the optimal contract.

PROPOSITION 2. *The IT value V increases monotonically in* $\bar{P}_i$, β_i ($i = 1, 2$), Q, *and* λ.

Proposition 2 suggests that the effects of exogenous project parameters Q and λ are in the expected direction. Note that Q and λ are independent of $\bar{P}_i$ or β_i. Consequently, the monotonicity of project value in general IT implementation capabilities ($\bar{P}_i$) or vendor expertise (β_i) does not depend on exogenous project parameters Q and λ. Therefore, we may characterize different learning conditions without being concerned that the model will yield implausible results for some values of the exogenous project parameters.

We now compare a multistage with a single-stage "big bang" contract. If we consider discounting, the optimal IT value for a single-stage (shown in §4.1) is: $V^*(\text{single-stage contract}) = V^*(1) = \lambda \Delta_1 - \pi^0$. Therefore, the difference of optimal project value between a two-stage contract and a single-stage contract is:

$$V^*(\text{two-stage contract}) - V^*(\text{single-stage contract})$$

$$= \lambda s_2 (\bar{P}_1 \Delta_2 - \Delta_1) + (1 - \lambda) s_1 \Delta_1 - \frac{s_1}{\beta_1} \ln\left(1 + \frac{\lambda s_2 \Delta_2}{s_1 Q}\right). \quad (8)$$

From (8) we see that the additional learning benefit due to staging is $(\lambda s_2 (\bar{P}_1 \Delta_2 - \Delta_1) + (1 - \lambda) s_1 \Delta_1)$, at the additional cost/effort of $(s_1/\beta_1) \ln(1 + \lambda s_2 \Delta_2 / (s_1 Q))$. We summarize this finding in Proposition 3:

PROPOSITION 3. *A two-stage contract is preferable to a single-stage contract if, and only if, the benefit of staging outweighs the cost of staging, i.e., if there exists* $\hat{\alpha}$, $0 < \underline{\alpha} \leq \hat{\alpha} < 1$, *such that* (8) > 0 *or equivalently,* $V^*(\hat{\alpha}) > V^*(1)$.

We illustrate the insights in Proposition 3 with two simple examples.

EXAMPLE 1. Assume $Q = 2$. If there is: no learning ($\beta_2 = \beta_1 = 1$), no capability improving ($\bar{P}_2 = \bar{P}_1 = 1$),

and no discounting ($\lambda = 1$), then (8) $= -(s_1/\beta_1) \ln(1 + \lambda s_2 \Delta_2 / (s_1 Q)) < 0$. A single-stage contract is preferable to a two-stage contract, because in this case staging incurs additional cost but gains nothing.

EXAMPLE 2. Retain other assumptions of Example 1 but allow for some agent learning due to staging such that $\beta_2 > e \approx 2.718 > 1 = \beta_1$. We have $\Delta_2 - \Delta_1 = (e - 2 - \ln 2)/e + \ln 2 > \ln 2$. If we set $s_1 = 1/2$, we have (8) $= 0.5(\Delta_2 - \Delta_1) - 0.5 \ln[1 + \Delta_2/Q] > 0.5 \ln 2 - 0.5 \ln[1 + \Delta_2/Q] > 0$. A two-stage contract is preferable to a single-stage contract, because the additional benefit due to staging offsets the additional cost of staging.

Generally speaking, a two-stage contract is preferable to a one-stage contract if there is sufficient learning benefit following a successful period 1 outcome such that $\bar{P}_1 \Delta_2 > \Delta_1$ (which can be satisfied if β_2 is sufficiently larger than β_1 and $\bar{P}_2$ is sufficiently larger than $\bar{P}_1$).

The above examples and discussion highlight the importance of learning and capability in two-stage contract design. We are interested in when a unique interior solution to the optimal staging problem exists. We begin with the most straightforward learning process, in which expertise increases with the length of the first stage. This yields the following assumption:

ASSUMPTION A1 (CONCAVE LEARNING). *Assume* β_1 *is fixed,* $\beta_2(\alpha)$ *is increasing and concave over* α.

Assumption A1, especially the concavity requirement, plausibly describes discovery learning processes in which firms learn additional details as a project progresses until they know essentially everything and can learn little more. Nearly all learning functions in the literature satisfy this condition (e.g., Lilien et al. 1992). We obtain the following general properties of learning and project sizing under optimal contracting:

THEOREM 1. *Assume A1. Assume both project revenue* Q *and initial expertise* β_1 *are sufficiently large so that the project is feasible. Let* $\hat{\alpha} \geq \underline{\alpha}$ *be the smallest project size that satisfies* $V^*(\hat{\alpha}) > V^*(1)$, *then a unique optimal sizing exists* ($\hat{\alpha} \leq \alpha^* < 1$); *otherwise a single-stage contract* ($\alpha^* = 1$) *is optimal.*

When the project revenue Q and initial expertise β_1 are sufficiently large, Theorem 1 shows that the optimal sizing of a project depends on vendor learning and capability increase. The IT payoff function is well behaved, in fact concave ("Inverted-U") over project size (see the proof of Theorem 1), so any hill climbing method will reach the optimal interior point. An approximation of the optimal sizing is given implicitly by $\alpha^* \approx 1 - [(\bar{P}_1 \lambda \Delta_2 - \Delta_1)/(\bar{P}_1 \lambda)](\partial \Delta_2 / \partial \alpha)^{-1}$.

To further explore the role of learning, it is useful to explicitly provide a learning function so that

we can characterize how the rate of learning affects project sizing. Consider a general class of functions that takes the form of $\beta_2(\alpha) = \beta_1 + (\bar{\beta}_2 - \beta_1)\alpha^{1/n}$, where n is interpreted as the rate of learning. This functional form accommodates both linear and power law relationships that have been discussed in the prior literature, and this functional form anticipates a function where additional project-specific experience enables an agent to become more capable, up to some upper limit ($\bar{\beta}_2$ is a constant that bounds β_2). When $n \geq 1$, this functional form yields a concave relationship between experience and expertise, consistent with Assumption A1. Further, when $0 < n < 1$, the form is also capable of capturing a convex relationship between experience and expertise. With this additional assumption we can show:
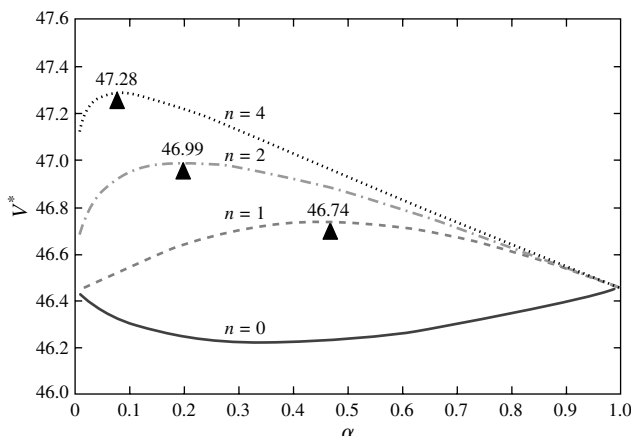
COROLLARY 1. *Assume* $\beta_2(\alpha) = \beta_1 + (\bar{\beta}_2 - \beta_1)\alpha^{1/n}$. *Then: the faster the rate of learning (larger n), (i) the larger the IT project value, i.e., $\partial V^*/\partial n > 0$, $\forall n > 0$, and (ii) the smaller the optimal first-period project size α^*, i.e., $d\alpha^*/dn < 0$, $\forall n \geq 1$.*

As illustrated in Figure 2, Corollary 1 provides the intuitive result that a faster learning rate always allows the client to capture more project value. Moreover, if a project is to be staged, faster learning leads to a smaller first stage. If we assume no discounting, the shape of the project value function depends on the speed of learning. Learning can change its shape from convex to concave. If we assume a moderate level of discounting, as we show later in Corollary 5, the concavity of the IT value function holds beyond Assumption A1 to include some cases of convex learning.

5. Extensions
Overall, Theorem 1 shows that as long as it is feasible to do a project, the optimal sizing depends on

Figure 2   Risk-Neutral Agent: Project Value, Learning and Optimal Staging

Note. $Q = 50$, $\bar{P}_1 = \bar{P}_2 = 1$, $r = 0$, $\pi_i^0 = 0$ $(i = 1, 2)$, $\beta_2 = \beta_1 + (\bar{\beta}_2 - \beta_1)\alpha^{1/n}$, $\beta_1 = 1.5$, $\bar{\beta}_2 = 3$, $\underline{\alpha} = 0.01$.

the learning rate, and that a faster learning rate yields higher project value yet a smaller first-stage size; a single-stage "big bang" project can be optimal under certain conditions (such as if there is no learning). In this section, we demonstrate that these general properties are quite robust to much broader assumptions than could reasonably arise in EIT projects.

5.1. Generalized Concave Learning
Several other plausible assumptions apply to the relationships among project sizing and learning. One possibility is that learning occurs both during the first and the second period $\beta_i(\alpha)$ $(i = 1, 2)$, i.e., β_1 is no longer fixed but is also a function of α; the conclusions of our preceding analyses generalize:

COROLLARY 2. *Suppose the assumptions used in Theorem 1 hold and a multistage contract is optimal. A multistage contract remains optimal if learning occurs both during the first and the second period, i.e., when $\beta_i(\alpha)$ $(i = 1, 2)$ are increasing and concave over α.*

5.2. Time-Varying Inherent Project Risk and Risk Reduction
In the first period, the inherent project risk may depend on the project size, such that a larger project may be more risky, and a "big bang" project has the highest inherent project risk $(1 - \bar{P}_1(\alpha = 1))$. It is equally plausible, after the successful completion of the first-period project, that the inherent project risk changes (i.e., decreases). These two generalizations can be summarized as follows.

ASSUMPTION A2 (LEARNING IN REDUCING INHERENT PROJECT RISKS). *$\bar{P}_1(\alpha)$ is decreasing and linear over α, and $\bar{P}_2(\alpha)$ is increasing and concave over α.*

This assumption is consistent with a view that a two-stage project can provide a Bayesian update of the risk of the overall project; the information updating of the second-period inherent risk $(1 - \bar{P}_2(\alpha))$ is conditional on first-period project success. Although our more general Assumption A2 is consistent with Bayesian updating, our interpretation hinges on the nature of the EIT project, in the sense that a successful first-period project removes some risks from the later stages, rather than that firms learn about a purely exogenous project risk.

COROLLARY 3. *Suppose the assumptions used in Theorem 1 hold and a multistage contract is optimal. A multistage contract remains optimal if there is learning that reduces inherent project risks (as described by A2).*

5.3. Learning Through Training
It is natural to believe that agents can acquire increased ability not only through project-specific experience, but also through project-specific training, which (in general) may be a function of project

structure since the training can occur during the first period. Let the training cost be represented by $t_1(\alpha)$. This gives rise to a more general concept of learning when $\beta_2(\alpha, t_1(\alpha))$ is a function of training. This cost could also be viewed as an investment by the agent in project-specific coordination. This can be formalized as follows:

ASSUMPTION A3 (LEARNING THROUGH TRAINING). $t_1(\alpha)$ *is increasing and linear over time. Furthermore,* $\beta_2(\alpha, t_1(\alpha))$ *is increasing and concave in training* $t_1(\alpha)$.

Intuitively, Theorem 1 holds under A3 because the introduction of an investment in training $t_1(\alpha)$ does not change the structure of our IT contract game, and the optimal contract design remains the same—except that we substitute agent reservation profit π_1^0 with $\pi_1^0 - t_1(\alpha)$.

COROLLARY 4. *Suppose the assumptions used in Theorem 1 hold and a multistage contract is optimal. A multistage contract remains optimal if there is learning through training (as described by* A3).

Collectively, these results suggest that the general structure of the results described by Theorem 1 is robust. Now we consider the impact of a continuous discounting rate.

5.4. Continuous Discounting

ASSUMPTION A4 (CONTINUOUS DISCOUNTING RATE). *Assume further the interest rate* $0 < r < 2$ *is the same for both the principal and the agent so that the discount factor at each period* i *is* e^{-rs_i} $(i = 1, 2)$.

COROLLARY 5. *Assume* A4. *Assume* Q *and* β_1 *are sufficiently large for the project to be feasible. Then* $\partial^2 V / \partial \alpha^2 < 0$.

Corollary 5 demonstrates the expected impact of discounting. If the principal and the agent are moderately patient—the interest rate is positive but below a reasonable threshold such that $r < 2$—then the IT payoff function becomes concave. Discounting at a reasonable interest rate tends to preserve or strengthen the concavity of the IT payoff function, so we extend Corollaries 1–4 to the case of continuous discounting, which suggests additional robustness of the structure of the results described by Theorem 1.

5.5. Risk-Averse Agents

We now consider the possibility that the agent is risk averse. For tractability, we consider risk-neutral principals and agents that show constant absolute risk aversion (CARA). CARA is commonly used in principal–agent modeling when the agent is risk averse (e.g., Holmstrom and Milgrom 1987).

ASSUMPTION A5 (CARA UTILITY). *The agent has a constant absolute risk aversion; that is,* $u(x) = -e^{-\gamma x}$ *for some* $\gamma \geq 0$.

To focus on the impact of the agent's risk tolerance (captured by γ), and without loss of generality, we let $\bar{P}_1 = \bar{P}_2 = 1$, $\pi_1^0 = \pi_2^0 = \pi^0 = 0$. We denote $u(b_2 R_2) \equiv w$, $u(b_1 R_1 + e^{-rs_2} \pi_2^0) \equiv v$, and $(s_i / \beta_i) \gamma \equiv B_i$, for $i = 1, 2$. We add a superscript A to our key parameters to denote risk attitude. By applying the *certainty equivalent principle*, we solve again the two-stage principal–agent problem. We summarize our findings in Theorem 2. In equilibrium, the optimal long-term contract $[a_i^A, b_i^A; s_i]$ $(i = 1, 2)$ and the agent's reaction becomes

$$a_1^A = \frac{1}{\gamma} \ln \frac{1+v}{B_1} + \frac{1-B_1}{\gamma} \ln \left[\frac{B_1}{1-B_1} \frac{-v}{1+v} \right] + \pi^0, \quad (9)$$

$$b_1^A = 1 + e^{-rs_2} \frac{\Pi_2^A}{R_1} + \frac{(1+v) + (1-B_1)(1+v^{-1})}{\gamma R_1}$$

$$< 1 + e^{-rs_2} \frac{\Pi_2^A}{R_1}, \quad (10)$$

where $\Pi_2^A = (1 - b_2^A) p_2^A R_2 - a_2^A$,

$$x_1^A = \frac{s_1}{\beta_1} \ln \left[\frac{1-B_1}{B_1} \frac{1+v}{-v} \right],$$

$$a_2^A = \frac{1}{\gamma} \ln \frac{1+w}{B_2} + \frac{1-B_2}{\gamma} \ln \left[\frac{B_2}{1-B_2} \frac{-w}{1+w} \right] + \pi_2^0, \quad (11)$$

$$b_2^A = 1 + \frac{(1+w) + (1-B_2)(1+w^{-1})}{\gamma R_2} < 1, \quad (12)$$

$$x_2^A = \frac{s_2}{\beta_2} \ln \left[\frac{1-B_2}{B_2} \frac{1+w}{-w} \right].$$

The IT project value is

$$V^A = e^{-r\alpha} \left(p_1^A \frac{(1+v-B_1)(1+v)}{-\gamma v} - a_1^A \right). \quad (13)$$

THEOREM 2. *Assume* A5. *Let* $\bar{P}_1 = \bar{P}_2 = 1$. *Given any sizing of an IT project* s_i $(i = 1, 2)$, *when a feasible contract exists, then it is specified in Equations* (9)–(12), *with project value given by* (13). *Further, as the agent becomes more risk averse (but not extremely risk averse), the principal needs to reduce the revenue-sharing factor and increase the fixed payment part. Formally,* $\partial b_i^A / \partial \gamma < 0$, $\partial a_i^A / \partial \gamma > 0$, $i = 1, 2$.
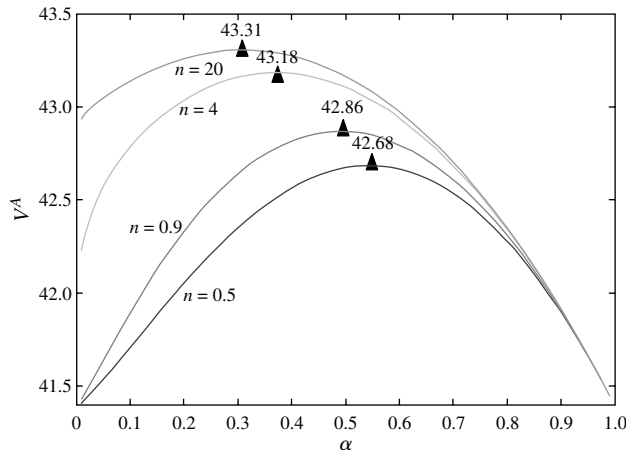
Note that, when initial expertise (β_1) is sufficiently large or when the agent's risk tolerance is not so large ($\gamma < \beta_1 < \beta_2$) then a feasible contract exists.

COROLLARY 6. *Assume* A1 *and* A5. *Let* $\bar{P}_1 = \bar{P}_2 = 1$. *Assume fixed discount factors. Assume both project revenue* Q *and initial expertise* β_1 *are sufficiently large. Assume a feasible contract exists, then the IT value function* V^A *is concave in* α. *Formally,* $\partial^2 V^A / \partial \alpha^2 < 0$.

We now characterize conditions when the optimal sizing is an interior point, and vice versa.

Define $\hat{\alpha}$ as the smallest α that satisfies $V^A(\hat{\alpha}) \geq V^A(1)$. Corollary 6 implies that if $\hat{\alpha}$ exists, then an

Figure 3   CARA Agent: Project Value, Learning and Optimal Staging

Note. $Q = 50$, $\bar{P}_1 = \bar{P}_2 = 1$, $\gamma = 0.5$, $r = 0.1$, $\pi_i^0 = 0$ ($i = 1, 2$), $\beta_2 = \beta_1 + (\bar{\beta}_2 - \beta_1)\alpha^{1/n}$, $\beta_1 = 1.5$, $\bar{\beta}_2 = 3$, $\underline{\alpha} = 0.01$.

interior optimal solution $\hat{\alpha} \le \alpha^* < 1$ exists, meaning a two-stage contract is optimal; otherwise a single-stage "big bang" contract is optimal.

Finally, we show that our key findings—agent learning increases client IT payoff when the agent is risk neutral—continue to hold when the agent is risk averse.

Corollary 7. *Assume* A1 *and* A5. *Let* $\bar{P}_1 = \bar{P}_2 = 1$. *Assume a general learning function such that* $\partial\beta_2/\partial n > 0$, $\forall n > 0$. *Then* $\partial V^A/\partial n > 0$.

As illustrated in Figure 3, key findings in the risk-neutral case are robust when the agent is CARA risk averse. Agent learning allows the client to capture more project value, and staging can be desirable even in the absence of learning, in contrast to the risk-neutral case where an absence of learning implies that a single-stage contract is preferable. As the agent becomes more risk averse, the principal scales back the revenue sharing factor but increases the fixed fee payment. Conversely, when the agent becomes less risk averse, the principal increases the incentives and reduces the fixed fee payment. Consistent with the literature in contract theory (Chiappori et al. 1994), it is possible that the individual rationality constraint cannot be satisfied, so there is no equilibrium with agent participation and no opportunity for a feasible contract. As with much of the principal–agent literature, we do not consider risk-averse principals due to the additional complexity, although this scenario could potentially be studied using simulation.

6.   Experimental Evidence

In this section, we focus on experiments related to the main predictions of our model in the context of EIT

contracting. We are interested in testing whether staging the project has the expected benefits under different learning rates. We are also interested in testing if a faster learning rate encourages decision makers to select a smaller first-stage size.

In our experiments, subjects act as consultants (representing the agents) playing against a computer (which represents the client or the principal) in a two-stage game in which the client contracts with the consultant for an IT project. A subject, presented with an effort-outcome mapping table (computed using our model), decides how much effort x to invest at each stage. After the investment, the binary outcome (success or failure) is realized and shown to the subject. We examine how experimental subjects perform relative to our theoretical predictions in Corollary 1.

We test three rates (slow, medium, fast) of agent learning. All treatments have the same expected profit and parameter set (except learning rate). Table 1 summarizes our experiments' design.

The risk-neutral model offers optimal contract parameter settings, which are conservative estimates compared with those produced by the risk-averse model (from Theorem 2). We parameterize our learning function in Corollary 1, $\beta_2 = \beta_1 + (\bar{\beta}_2 - \beta_1)\alpha^{1/n}$ with $\bar{\beta}_2 = 1$ and three scenarios of learning: $n = 0.5$ for slow learning (Treatment 1, T_1), $n = 0.9$ for medium learning (Treatments 2 and 3, T_2 and T_3), and $n = 20$ for fast learning (Treatment 4, T_4). We test how learning allows the client to receive higher value by staging the project. Specifically, our model predicts the following:

Hypothesis 1 (H1). *Client project value increases* (a) *from* T_1 *to* T_3, (b) *from* T_3 *to* T_4, *and* (c) *from* T_1 *to* T_4.

To test whether faster learning reduces first-period project size, controlling for client profit, we added treatment T_2 which has a smaller first-period size compared with T_1. We examine this prediction indirectly through the consultant effort level at Stage 1, which should decrease due to size reductions; hypothesized as follows:

Hypothesis 2 (H2). *First-period consultant effort* x_1 *decreases from* T_1 *to* T_2.

We conducted three series of experiments with different subjects for a total of eight sessions lasting 90 minutes each. In Experiment 1, our baseline

Table 1   Treatments

	Slow learning ($n = 0.5$)	Medium learning ($n = 0.9$)	Fast learning ($n = 20$)
Two-stage incentive payment (suboptimal)		T_2	
Two-stage incentive payment (optimal)	T_1	T_3	T_4

Table 2   Parameterization for Real Run Experiments

	External environment				Stage 1 contract		Stage 2 contract		Client profit
	α	β_1	n	β_2	a_1	b_1	a_2	b_2	V
T_1	0.3	0.1	0.5	0.181	−22.547	1.894	−26.812	1.000	22.547
T_2	0.128	0.1	0.9	0.192	−22.547	3.822	−36.178	1.000	22.547
T_3	0.3	0.1	0.9	0.336	−25.845	2.061	−31.842	1.000	25.845
T_4	0.3	0.1	20.0	0.947	−29.020	2.222	−36.662	1.000	29.020

Notes. Across all treatments, $Q = 100$, $\bar{P}_1 = \bar{P}_2 = 0.7$, $r = 0$. The expected consultant profit remains the same, $\pi_i^0 = 12$ ($i = 1, 2$), across all treatments.

experiment, we recruited 7 subjects each in Treatments 1–4 from a general student pool (undergraduate students randomly recruited on campus via campus advertisements) for two sessions.[8] We conducted two additional series of experiments for robustness in another location. Experiment 2 used 9, 10, 16, and 13 subjects, respectively, in Treatments 1–4. Subjects in Experiment 2 were graduate and undergraduate students who had completed half of a semester-long course on IT management and were familiar with topics such as ERP systems, implementation, and outsourcing. Finally, the single session in Experiment 3 involved 8 and 7 executives in Treatments 1 and 4, respectively; their average 17 years of industry experience makes Experiment 3 essentially a controlled field experiment (e.g., Harrison and List 2004).

To provide corresponding incentives to the subjects, we used cash payments for Experiment 1, course credits for Experiment 2, and lottery money for Experiment 3. Subjects could "test drive" the system during 10 dry runs before they participated in the 30 real runs. Although the structure of the game remained the same between dry and real runs, we used slightly different parameters.[9] We summarize the parameterization of our real run experiments in Table 2.

We summarize our findings in Tables 3–5. We find that, overall, the client is able to capture more project value from increased consultant skill via learning and staging, as H1(b) and H1(c) are both supported. In Experiment 1, client profit increases from 17.247 in T_1 to 18.486 in T_3 ($t = 0.941$, n.s.) to 21.466 in T_4 ($t = 2.637$, $p < 0.1$). Results are similar in Experiment 2 and Experiment 3 (when applicable). We find only partial support for H1(a), as client profit increase from T_1 and T_3 in Experiment 1 is not significant, perhaps due to small changes in learning rates (from $n = 0.5$ to $n = 0.9$); it turns out to be significant in Experiment 2, as

client profit increases from 16.625 in T_1 to 18.218 in T_3 ($t = 2.01$, $p < 0.1$).

Our analysis also provides partial support for H2. In Experiment 1, consultants scale back their Stage 1 effort from 10.257 (T_1) to 5.586 (T_2) in response to a reduced first-stage project size ($\alpha = 0.3$ in T_1 versus $\alpha = 0.128$ in T_2), which is significant ($t = 4.869$, $p < 0.001$). Similarly, in Experiment 2, consultants reduce their Stage 1 effort from 13.966 (T_1) to 10.937 (T_2), though the change is not significant ($t = 1.042$, n.s.). Support for H2 suggests indirectly that the client can reduce IT project risk by awarding a smaller, rather than larger, first-stage project to a faster learning team.

In summary, our experiments reveal that learning and staging in contracting play an important role in creating and capturing IT value for the client. Interestingly, student subjects tend to expend more effort than necessary, while the executive subjects seem to be more rational in choosing the theoretically optimal effort levels. This suggests potential future extensions of our model to explore the heterogeneity in IT expertise across subject groups.[10] Overall, our findings are robust across experiments; and we did not expect and did not find significant qualitative behavior differences among subjects. An analysis of click-through data also shows no significant differences across treatments, subjects, or stages.[11]

7.   Conclusions

By constructing a principal-agent model and validating our model experimentally, we study how firms can create and capture EIT project value via IT contract design. Our model integrates three previously isolated streams of research—dynamic production functions, multiperiod moral hazard, and IT outsourcing and contracting—in the context of EIT project management. Specifically, we create a single model that links salient features of EIT markets (contract choice, learning, risk management practice) to project outcome.

---

[8] Total numbers of subjects in each experiment were 42 (Experiment 1), 73 (Experiment 2) and 24 (Experiment 3), respectively. The remaining subjects were used for additional treatments (not reported here for brevity, but available from the authors upon request) for further robustness testing of our modeling predictions.

[9] Our design of slightly different parameterizations for the dry versus real runs follows the experimental economics literature and ensures the attention of the subjects.

[10] We thank a referee for bringing this opportunity to our attention.

[11] Our click-through data analysis is available upon request; we omit it here for brevity.

Table 3   Summary of Predicted versus Observed Consultant Efforts (Mean and Variance)

| | Stage 1 effort (x_1) | | | | Stage 2 effort (x_2) | | | |
| | | Observed | | | | Observed | | |
	Predicted	Exp. 1	Exp. 2	Exp. 3	Predicted	Exp. 1	Exp. 2	Exp. 3
T_1	8.172	10.257 (3.290)	13.966 (28.262)	8.583 (4.545)	9.820	12.775 (4.448)	16.850 (35.652)	10.692 (2.654)
T_2	4.418	5.586 (3.153)	10.937 (53.134)	N/A	11.805	12.392 (7.516)	15.109 (66.812)	N/A

Note. The mean is shown at the top of each cell; variance is shown in parenthesis.

Table 4   Summary of Predicted versus Realized Client/Consultant Profits

| | Client profit | | | | Consultant profit | | | |
| | | Realized | | | | Realized | | |
	Predicted	Exp. 1	Exp. 2	Exp. 3	Predicted	Exp. 1	Exp. 2	Exp. 3
T_1	22.547	17.247	16.625	16.695	12	12.352	5.650	10.459
T_2	22.547	13.952	13.476	N/A	12	12.743	4.111	N/A
T_3	25.845	18.486	18.218	N/A	12	5.486	6.609	N/A
T_4	29.020	21.466	21.335	20.652	12	−0.642	6.508	4.887

Theoretically, we extend previous multiperiod moral hazard models to incorporate (1) learning and dynamic production functions, (2) risk diversification and reduction over project periods, and (3) linkages (contingent contracts) between periods. These characteristics are particularly salient to large-scale EIT projects but have not been formally treated or experimentally tested in the literature.

In turn, our analysis yields several new insights into contract design. First, we examine the role of agent learning. In general, multistage contracts tend to be favorable in the presence of vendor learning or the possibility of exogenous risk reduction. Gains from staging depend on the rate of learning and can be large–relative to project revenue. Our results are robust across situations in which learning depends not just on first-period project size but also on investments in training. In the absence of learning and discounting, single-stage "big bang" projects are generally preferable over staged projects among risk-neutral agents. Our results appear to grow stronger when we consider risk-averse agents. Finally, we provide a theoretical

basis for the conditions in which EIT value creation takes an inverted-U form, and thus provide theoretical underpinnings for recent empirical findings about EIT value creation associated with EIT project management (e.g., Aral et al. 2006, Bouhdary and Comes 2008, Hitt et al. 2002).

Using controlled lab experiments, we test theory-based relationships among IT project value, learning, and staging in contracting. We show that learning allows the client to receive higher value by staging the project, and a faster learning rate results in a smaller first-stage project size. Experiments suggest that our theoretical results are robust to deviations from our specific assumptions about production functions or agent risk attitudes. Academically, these findings extend the literature on software contracting from development (e.g., Lee and Png 1990, Wang et al. 1997, Whang 1992) to off-the-shelf software service (e.g., implementation and maintenance). Practically, they have immediate implications for IT contract design in EIT markets and also in the recently emerged SaaS (software-as-a-service) markets where traditional EIT software vendors are in transition (Chou 2010), suggesting significant benefits from relatively small changes to contract structure terms.

Our model and experiments highlight the importance of agent learning, patience, and risk attitude, and their relationship with optimal project staging, in the context of EIT project risk management. Although our results provide some initial insights into one aspect of contracting for enterprise software projects, opportunities remain for extending these models and empirically investigating the relationship between EIT

Table 5   Summary of Results (t-statistics)

Hypotheses	Experiment 1	Experiment 2	Experiment 3
H1(a): V increases from T_1 to T_3	0.941	2.010*	N/A
H1(b): V increases from T_3 to T_4	2.637*	4.375***	N/A
H1(c): V increases from T_1 to T_4	5.614***	6.046***	5.205***
H2: x_1 decreases from T_1 to T_2	4.869***	1.042	N/A

Notes. One-tailed t-test assuming two-sample equal variances for Experiment 1 and unequal variances for Experiments 2 and 3.
   ***$p < 0.001$, **$p < 0.01$, and *$p < 0.1$.

value creation and contract structure using laboratory or natural experiments. Some potential extensions of our model include the following: examining the optimal contracting across repeated projects where there may be learning across projects; modeling contracting for simultaneous projects, in the presence of agent learning; examining the impact of competition among vendors; or modeling the diseconomies of splitting a project. We anticipate that a fruitful line of research can be built upon our model.

Supplemental Material

Supplemental material to this paper is available at http://dx .doi.org/10.1287/isre.1120.0448.

Acknowledgments

References

Anderson EG (2001) Managing the impact of high market growth and learning on knowledge worker productivity and service quality. *Eur. J. Oper. Res.* 134(3):508–524.

Aral S, Brynjolfsson E, Wu DJ (2006) Which came first, IT or productivity? The "virtuous cycle" of investment and use in extended enterprise systems. *Proc. 27th Internat. Conf. Inform. Systems, Milwaukee, WI.*

Banker R, Davis G, Slaughter S (1998) Software development practices, software complexity, and software maintenance. *Management Sci.* 44(4):433–450.

Barry E, Mukhopadhyay T, Slaughter S (2002) Software project duration and effort: An empirical study. *Inform. Tech. Management* 3(1):113–136.

Bolton P, Dewatripont M (2005) *Contract Theory* (MIT Press, Cambridge, MA).

Bouhdary C, Comes S (2008) Accelerate value creation: The virtuous cycle of using technology to maximize business value. *SAP Executive Insight* (October), 1–8.

Brooks FP (1975) *The Mythical Man Month* (Addison Wesley, Reading, MA).

Brynjolfsson E, Fitoussi D, Hitt L (2006) The information technology iceberg. Working paper, MIT Sloan School of Management, Cambridge, MA.

Chellappa R, Saraf N (2010) Alliances, rivalry and firm performance in enterprise systems software markets: A social network approach. *Inform. Systems Res.* 21(4):849–871.

Chiappori PA, Ines M, Patrick R, Salanie B (1994) Repeated moral hazard: The role of memory, commitment, and the access to credit markets. *Eur. Econom. Rev.* 38(8):1527–1553.

Chou T (2010) *Cloud: Seven Clear Business Models* (Active Book Press, Los Altos Hills, CA).

Choudhury V, Sabherwal R (2003) Portfolios of control in outsourced software development projects. *Inform. Systems Res.* 14(3):291–314.

Clemons E, Hitt L, Snir E (2001) A risk analysis framework for IT outsourcing. Wharton OPIM Working paper, The University of Pennsylvania, Wharton School, Philadelphia.

Cohen MC, Agrawal N, Agrawal V (2006) Winning in the aftermarket. *Harvard Bus. Rev.* 84(5):129–138.

Davenport TH (2005) The coming commoditization of processes. *Harvard Bus. Rev.* 83(6):101–108.

DiRomualdo A, Gurbaxani V (1998) Strategic intent for IT outsourcing. *Sloan Management Rev.* 39(4):67–80.

Gaimon C (1997) Planning information technology—knowledge worker systems. *Management Sci.* 43(9):1308–1328.

Gaimon C, Ozkan G, Napoleon K (2011) Dynamic resource capabilities: Managing workforce knowledge with a technology upgrade. *Organ. Sci.* 22(6):1560–1578.

Haines M, Goodhue D (2003) Implementation partner involvement and knowledge transfer in the context of ERP implementations. *Internat. J. Human-Comput. Interaction* 16(1):23–38.

Harrison G, List J (2004) Field experiments. *J. Econom. Literature* 42(4):1009–1055.

Hitt LM, Wu DJ, Zhou X (2002) ERP investment: Business impact and productivity measures. *J. Management Inform. Systems* 19(1):71–98.

Holmstrom B (1982) Moral hazard in teams. *The Bell J. Econom.* 13(2):324–340.

Holmstrom B, Milgrom P (1987) Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55(2): 303–328.

James G (1998) Method of payment. *CIO Enterprise* (October 15, Section 2):92–95.

Kirsch LJ (2000) Software project management: An integrated perspective for an emerging paradigm. Zmud RW, ed. *Framing the Domains of IT Management Research: Glimpsing the Future Through the Past*, Chap. 15. (Pinnaflex Educational Resources Inc., Cincinnati, OH), 285–304.

Lacity M, Hirschheim R (1993) The information systems outsourcing bandwagon: Look before you leap. *Sloan Management Rev.* 35(1):72–86.

Lacity M, Willcocks L (1998) Practices in information technology outsourcing: Lessons from experience. *MIS Quart.* 22(3): 363–408.

Lambert R (1983) Long-term contracts and moral hazard. *Bell J. Econom.* 14(2):441–452.

Lee TK, Png IPL (1990) The role of installment payments in contracts for services. *RAND J. Econom.* 21(1):83–99.

Levy F (1965) Adaptation in the production process. *Management Sci.* 11(6):B136–B154.

Lilien GL, Kotler P, Moorthy KS (1992) *Marketing Models* (Prentice Hall, Upper Saddle River, NJ).

Loch CH, Kavadias S (2002) Dynamic portfolio selection of NPD programs using marginal returns. *Management Sci.* 48(10):1227–1241.

Markus LM, Tanis C, van Fenema PC (2000) Multisite ERP implementations. *Comm. ACM* 43(4):42–46.

McAfee A (2002) The impact of enterprise information technology adoption on operational performance: An empirical investigation. *Production Oper. Management* 11(1):33–53.

McAfee A (2003) When too much IT knowledge is a dangerous thing. *MIT Sloan Management Rev.* 44(2):83–89.

McAfee RP, McMillan J (1986) Bidding for contracts: A principal–agent analysis. *RAND J. Econom.* 17(3):326–338.

McAfee RP, McMillan J (1987) Competition for agency contracts. *RAND J. Econom.* 18(2):296–307.

McFarlan FW (1981) Portfolio approach to information systems. *Harvard Bus. Rev.* 59(5):142–150.

Mendelson H (2000) FoxMeyer's Delta project. Stanford Graduate School of Business Case, Stanford University, Stanford, CA.

Nolan R, McFarlan FW (2005) Information technology and the board of directors. *Harvard Bus. Rev.* 83(10):96–106.

O'Leary D (2002) *Enterprise Resource Planning Systems: Systems, Life Cycle, Electronic Commerce, and Risk* (Cambridge University Press, New York).

Pike G (2006) Supporting business innovation while reducing technology risk. White paper, SAP AG, Walldorf, Germany.

Retrieved September 21, 2012, http://www.sap.com/malaysia/solutions/safeguarding/pdf/supportBiz20sep.pdf.

Radner R (1981) Monitoring cooperative agreements in a repeated principal–agent relationship. *Econometrica* 49(5):1127–1148.

Radner R (1985) Repeated principal–agent games with discounting. *Econometrica* 53(5):1173–1198.

Richmond WB, Seidmann A (1993) Software development outsourcing: Contract structure and business value. *J. Management Inform. Systems* 10(1):57–72.

Richmond WB, Seidmann A, Whinston AB (1992) Incomplete contracting issues in information systems development outsourcing. *Decision Support Systems* 8(4):459–477.

Rogerson WP (1985) Repeated moral hazard. *Econometrica* 53(1):69–76.

SAP annual report (2008) Retrieved December 21, 2010, http://www.sap.com/about/investor/reports/annualreport/2008/pdf/SAP_2008_Annual_Report.pdf.

Scott J (1999) The FoxMeyer Drugs' bankruptcy: Was it a failure of ERP? *Proc. Fifth Americas Conf. Inform. Systems, Milwaukee, WI,* 223–225.

Snir E, Hitt L (2004) Vendor screening in information technology contracting with a pilot project. *J. Organ. Comput. Electronic Commerce* 14(1):61–88.

Umble EJ, Haft RR, Umble MM (2003) Enterprise resource planning: Implementation procedures and critical success factors. *Eur. J. Oper. Res.* 146(2):241–257.

Wang ETG, Barron T, Seidmann A (1997) Contracting structures for custom software development: The impact of informational rents and uncertainty on internal development and outsourcing. *Management Sci.* 43(12):1726–1744.

Whang S (1992) Contracting for software development. *Management Sci.* 38(3):307–324.