

PRB
7/17/08
11:15 am

NBER 2008 Summer Institute
Productivity Potpourri Workshop

Thursday, July 17, 2008

Organizer
Ernst R. Berndt (MIT and NBER)

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

D.J. Wu
College of Management, Georgia Institute of Technology, Atlanta, GA 30332,
dj.wu@mgt.gatech.edu

Min Ding
Smeal College of Business, Pennsylvania State University, University Park, PA 16802,
minding@psu.edu

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

Discussant
Shane Greenstein (Northwestern and NBER)

IT Implementation Contract Design: Analytical and Experimental Investigation of IT Value, Learning, and Contract Structure¹

D.J. Wu

College of Management, Georgia Institute of Technology, Atlanta, GA 30332,
dj.wu@mgt.gatech.edu

Min Ding

Smeal College of Business, Pennsylvania State University, University Park, PA 16802,
minding@psu.edu

Lorin M. Hitt

The Wharton School, University of Pennsylvania, Philadelphia, PA 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, which they use to develop an analytical model of optimal contract design with principal-agent techniques. The model captures 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 shows that incentive contracts with a multiperiod structure, combined with the appropriate project scope, create greater overall value, because they prompt the implementation team to exert greater effort while also capturing learning benefits and reducing risk. The results are consistent with prior empirical findings that payoffs from enterprise resource planning implementation are concave with the scope of the project. To test and validate the model predictions and key assumptions, the authors use controlled laboratory experiments.

Key words: enterprise information technology; enterprise systems; IT contracting; principal-agent; multi-period

¹ Draft for comments. Please do not circulate without permission of the authors. Please send comments to dj.wu@mgt.gatech.edu.

1. Motivation

Software as a percentage of total private fixed investment has exceeded computer hardware in the United States since 1992,² driving the U.S. corporate information technology (IT) stock per full-time employee up from \$779 in 1987 to \$2,646 in 2004 (Figure 1a; Brynjolfsson et al. 2006b). Enterprise information technology (EIT), such as enterprise resource planning (ERP), supply chain management (SCM), and customer relationship management (CRM) systems, represents the largest single category of IT spending (Brynjolfsson et al. 2006a), accounting for more than \$40 billion in annual investments (Aral et al. 2006). These systems furthermore serve as catalysts for the order-of-magnitude greater investments in organizational capital (e.g., process redesign, training, configuration, deployment), which may represent a capital stock of \$1.82 trillion, more than 15 times the market price of U.S. gold reserves at Fort Knox and a major share of all capital in the U.S. economy (Brynjolfsson et al. 2006a).

Large-scale packaged software systems typically get outsourced to consulting companies for implementation and maintenance. Such projects are notoriously complicated and place increasing burdens on the design of IT outsourcing relationships, especially 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 atypical for these projects to include various contracting opportunities over multiple years (e.g., O'Leary 2002). Despite these investments, evidence indicates that outsourcing agreements often perform less than satisfactorily (e.g., Lacity and Hirscheim 1993, Lacity and Willcocks 1998).

Contracts provide the primary means of IT governance (Nolan and McFarlan 2005), especially for codifying client–vendor relationships in IT outsourcing agreements (Clemons et al. 2000). 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 in-

² According to the fixed asset tables provided by the U.S. Bureau of Economic Analysis (2007), from 1992 to 2002, computers accounted for a steady 4% to 6% of capital investment, but software investments as a proportion of all capital investment grew from approximately 7% to 12%.

dicates that many outsourcing agreements prove difficult to manage, as reflected in the serious problems associated with some major ERP implementations (e.g., McAfee 2003, 2006, Mendelson 2000).

Enterprise information technology implementation contracts tend to be similar across different projects because they rely on the commoditization and standardization of business processes (Brynjolfsson et al. 2006b, Davenport 2005). In addition, the industry consists of a few dominant software vendors and implementation consultants that generally share common business practices.³ These characteristics create a broad typology of enterprise software contracting and implementation practices. Projects also contain features common to regular IT contracts, though on a larger scale, such as the extensive use of outside consultants, the integration of packaged (or modified “off-the-shelf”) software, and largely observable initial results (i.e., the client can either “go live” with the system or not).

Our analysis focuses on additional features that, though often present in regular IT contracts, are particularly important for EIT implementations. For example, projects often involve multiple, discrete periods (or stages) during which similar activities get repeated at different work sites or for different subsystems (e.g., component modules, suites) in the course of a multiyear project. Thus, the project may be structured as a single-period “big bang,” in which the firm commits to a full implementation without any scheduled intermediate decision points, or a staged, multiperiod, incremental rollout. A staged contract enables the effective use of information gathered during the project, such as midterm project evaluations of unexpected risks that may lead to the early termination of unsuccessful projects.⁴ This approach also can provide stronger incentives to the vendor by establishing an additional outcome measure, even if project success or failure at each stage represents the only contractible outcome. Finally, this evaluation offers a means to assess best practices and disseminate them throughout the project team, so teams can im-

³ In 2005, shares in the enterprise market were as follows: SAP 62%, Oracle 25%, and Microsoft 13%, based on worldwide software license revenue (SAP 2005 annual report, financial summary). Some large implementation firms include Accenture, PriceWaterhouseCoopers, and BearingPoint.

⁴ FoxMeyer, the fourth largest U.S. drug distributor (with a \$5.2 billion revenue) filed for bankruptcy largely because of the failure of a time-pressured big bang ERP implementation that cost more than \$100 million. The firm sued both the software vendor and the implementation provider for lacking experience and using the client as a “guinea pig” to train consultants onsite (Mendelson 2000). McKesson, a FoxMeyer rival, successfully implemented its ERP around the same time using a phased approach, during which it changed the lead systems integrator twice for poor performance and scaled down the project size when the project ran into trouble (Mendelson 2000). McKesson eventually bought FoxMeyer.

prove 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. 2006a), both learning and setting the proper project scope for each incremental phase can have significant impacts on project value (e.g., Hitt et al. 2002, McAfee 2003). Consistent with these observations, Hitt et al. (2002) identify a concave relationship between the extent of implementation (number of modules) and business value, such that medium-scope implementations outperform larger and smaller ones. In addition, ERP vendors understand the risk management benefits of a flexible project scope:

Implementation of SAP ... often involves a significant commitment of resources by SAP's customers and is subject to a number of significant risks over which it has little or no control.... SAP customers nowadays increasingly follow modular project approaches to optimize their IT environment, that is, 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 projects to implement an all-embracing IT landscape. (SAP 2005 annual report, 84–85)

However, subdividing a project is not without costs, which may explain the use of big bang projects in practice. Breaking a project into stages may delay the implementation of later stages, which can reduce or delay realization of value. Staged implementation also may lead to overinvestment during early periods, because advancing to the next period requires success in the previous period, or offset learning benefits when the potential expertise increases in subsequent periods diminish.

To address these issues, this analysis theoretically models and experimentally tests the relationship among project value, learning, and contracting (scope and fee structure) in a multiperiod setting. All else being equal, firms that adopt multiperiod incentive contracts with an appropriate project scope should create greater overall IT value, because they prompt the implementation team to exert greater effort, capture learning benefits, and mitigate risks. However, in some conditions, a single-stage contract structure can be optimal.

Our analytical model relates project value to vendor learning and contract structure and thus captures salient features we observe in our review of actual EIT outsourcing contracts. We also consider extensions to the model to allow for different structures of vendor learning. We validate the assumptions and key predictions of our model using a series of laboratory experiments, which largely confirm the ben-

efits of a multistage contract structure, especially in terms of the ability to elicit greater vendor effort through incentives and the potential to capture learning benefits. All proofs appear in the online Appendix A. Details of the laboratory experiments appear in the online Appendix B.

2. Literature Review

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

2.1. Learning and Dynamic Production Function

The first stream of related literature involves “learning by doing,” which characterizes 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 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. Her analysis suggests that an exponential functional form that incorporates a learning process meets all the required criteria and appears superior to some standard alternatives (e.g., Cobb-Douglas production; see also the survey by Gulledge and Khoshnevis 1987) 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 the 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 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 experience, effort, and project size in determining project outcomes, given the presence of controllable and uncontrollable project risks. Although IT implementation projects are risky by nature (McFarlan 1981, O’Leary

2002), empirical evidence in an ERP context suggests that managers and investors perceive particularly high risks (Hitt et al. 2002). 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. 2007). Thus, we extend Levy's model by allowing firms to reduce their project risks through effort and learning, but the project also suffers inherent project risks that must be managed but cannot be reduced by vendor effort.

2.2. Repeated 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-stage principal-agent problem (or parallels between the single stage and the infinitely repeated problem; see Holmstrom and Milgrom 1987). However, we know significantly less about finite-horizon repeated moral hazard games (e.g., Lambert 1983, Rogerson 1985), the setting that seems naturally to characterize large-scale IT implementations. One exception is Lambert (1983), who considers the role of long-term contracts for controlling two-period moral hazard problems and argues that the optimal contract in a finite horizon model is more complicated than those derived from infinite horizon models (e.g., Radner 1981, 1985), which parallel the same single-period game.

Our approach builds on these analyses but differs in several ways. First, we consider linkages (contingent contracts, revenue streams, and principal hold-up by the agent) between periods that are ignored in the repeated principal-agent literature (where identical single-period games are repeated finitely or infinitely) but essential to EIT projects. Second, we consider the potential for change in the effort-output relationship through agent learning and the uncertainty reduction achieved by updating information over the course of a project. Third, we develop our model specifically to capture key project characteristics observed from actual

IT contracts (e.g., project structure, payments). Finally, we test our model assumptions and key predictions experimentally.

2.3. IT Outsourcing and Contracting

The final stream of related literature pertains to outsourcing risk management (e.g., Clemons et al. 2000) 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 Hirsheim 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 updating information over the course of a project (Kulatilaka and Lin 2006, Snir and Hitt 2004, Whang 1992).

Through analyses of actual ERP contracts, we identify at least two major structural dimensions.⁵ First, though firms purchase ERP software externally, they must decide whether to perform the implementation in-house or contract with a service provider. Second, these contracts can be grouped into single-stage, big bang implementations or series of sequentially interlinked projects that give the client a termination decision at the end of each stage. Beyond these major decisions, the contracts principally differ in their use of fixed-fee versus incentive payments⁶ and the extent to which the vendor makes relationship-specific investments (e.g., training, facilities). 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.

⁵ Over a period of several years (2000–2006), students of two major U.S. universities searched and analyzed publicly available ERP contracts for course credit, drawing on sources provided by clients, ERP consultants, and software vendors. The research team also conducted site visits to examine real-world EIT projects and a set of real-life contracts.

⁶ Variable portions of payment 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 the post-ERP implementation. These structures are equivalent to a fixed-fee plus performance-based variable incentive structure, at least to the extent that their objective is to provide incentives to the vendor, and are consistent with incentives summarized in business 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).

Together, these arguments naturally suggest considering enterprise software contracting in a multiperiod setting, such that 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.

3. Model, Solution, and Discussions

3.1. Model

For ease of exposition and the experimental implementation, we restrict our model to a two-period setting in which two risk-neutral firms (client or principal and vendor or agent) face a principal–agent problem. We relax some of these assumptions in Section 5, but in the first analysis, we assume the principal can divide an enterprise IT project into two sequential periods (in a long-term contract), such that the agent is committed to complete the second stage if the principal desires to continue. Principal commitment is not required because it arises endogenously, as we discuss later. The initial participation of the agent is voluntary, so the agent must receive a guarantee to earn at least its reservation utility in both periods. The notation for our model appears in Table 1, and the timing and extensive form of the game are in Figure 1.

At the beginning of Period 1 (or time 0), the principal offers the agent a two-period contract in the form of $\mathcal{C} \equiv [a_i, b_i, s_i, i = 1, 2]$, where $s_1 \equiv \alpha$ denotes the size or scope of the project for the first period and $\underline{\alpha}$ denotes the smallest feasible project size (e.g., module). We normalize the overall project size to 1, so we can interpret α as project duration or time ($0 < \underline{\alpha} \leq \alpha \leq 1$). It is sometimes more intuitive to think about a project as two stages; therefore, we write the size of each stage as $s_1 = \alpha$ and $s_2 = 1 - \alpha$ (we use the α and s_i notation interchangeably). In practice, α represents the principal's choice of one 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). This subdivision also might apply at the module (e.g., Hitt et al. 2002) or suite (Aral et al. 2006) level, as well as in the form of a staged rollout of similar modules/suites across multiple

sites. For simplicity, we assume a continuous α .⁷ For each period $i = 1, 2$, a_i is a fixed fee, and b_i is a scale factor of an incentive payment $b_i R_i$, contingent on project success, where $R_i \equiv s_i Q$ is period i 's revenue or cash flow from a successful EIT project. Because we vary the parameters, our contract form $\mathcal{C} \equiv [a_i, b_i, s_i, i = 1, 2]$ is very general and captures a variety of practical project structures and fee structures, including a two-period phased approach (when $0 < \alpha < 1$) or one-period big bang approach ($\alpha = 1$); both fixed fee ($b_1 = b_2 = 0$) and incentive ($b_i > 0, i = 1, 2$) payments; and contracts with ($a_i \leq 0, i = 1, 2$) or without ($a_i > 0, i = 1, 2$) an initial financial commitment by the vendor. We allow for hybrid contracts that vary in structure by period (e.g., one-period fixed fee, two-period incentive payments), which reveals the optimal contract design within this general class of contract structures.

In each single period $i = 1, 2$, the agent exerts effort or action (and experiences cost) x_i , which is a real number and affects the probability of project success p_i at the end of each period. We assume the outcome of each period is a random variable that takes values of 1 (success) or 0 (failure).⁸ At the end of each period $i = 1, 2$, if the project fails, the agent receives its contractual payments, less its costs ($a_i - x_i$), and the principal receives ($-a_i$); if the project succeeds, the agent receives a full contractual payment ($a_i + b_i R_i - x_i$), and the principal receives full benefits less its contract costs ($R_i - a_i - b_i R_i$).

We assume the probability of project success p_i in each period may be influenced by two types of risks: inherent project risks \bar{P}_i , which capture technical, business, or managerial challenges that are part of the project (e.g., O'Leary 2002), and controllable risks, which are influenced by agent effort x_i , expertise (ability, competency, or "type") β_i , and project size s_i . We adopt conventional assumptions from contract

⁷ Requiring a discrete α does not change our findings qualitatively but would unnecessarily complicate the analysis. The value function we derive is typically concave with project size, so restricting the analysis to discrete sizes simply requires selecting the highest value of the two closest feasible points near the continuous optimum. When the value function is not concave, the optimum lies on the boundary, which is well defined regardless of whether we consider discrete or continuous sizing.

⁸We assume the EIT project is a success if the enterprise system can be turned on within the projected time period but a failure otherwise. Although other notions of success or failure may be applicable (e.g., smooth transitions, level of documentation), they

theory to assert that the principal and agent each can act as a single economic actor (thus separating any internal principal–agent problem within the client or vendor firm from the principal–agent problem between the two firms).⁹ To formalize controllable risks, we adopt a specific (dynamic) production function, namely, the exponential form $(1 - e^{-\beta_i \frac{x_i}{s_i}})$, to enable closed-form solutions and facilitate experimental implementation. Gaimon (1997) and Levy (1965) adopt similar production functions, perhaps because they represent one of the simplest functional forms that still capture the salient properties of learning-by-doing in IT contracting. Because controllable and uncontrollable risks jointly determine the probability of success and are statistically independent, we multiply the two types to obtain our overall project risk p_i . We thus reinterpret Levy’s workforce skill–outcome model for a setting in which skills and learning reduce risks to a minimum by overcoming controllable project risks.¹⁰

$$\text{ASSUMPTION (A1: Production Function). } p_1(x_1, s_1, \bar{P}_1, \beta_1) = \bar{P}_1(1 - e^{-\beta_1 \frac{x_1}{s_1}}), \text{ and}$$

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

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 type are known to the agent, the potential adverse selection issues (Snir and Hitt 2004, Whang 1992) presumably have been resolved, and our mod-

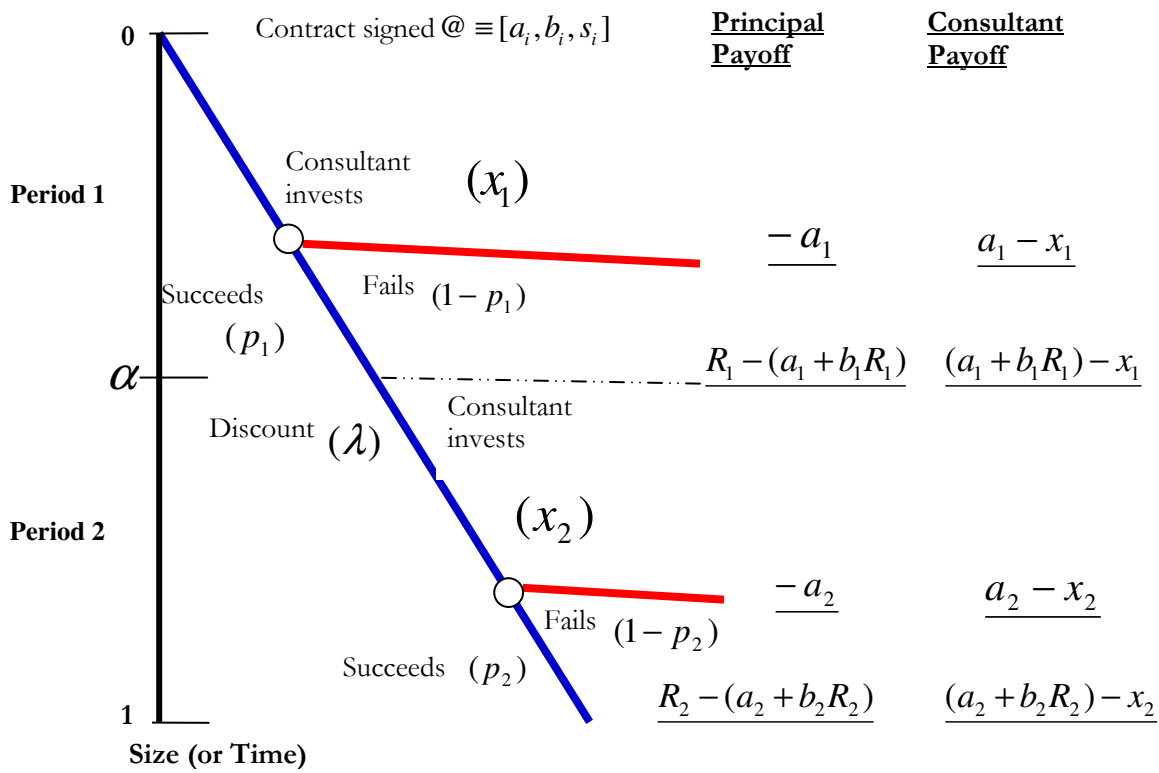
are harder to observe and contract. Moreover, a client can demand that these intangibles be incorporated into the project before going live, so our model can subsume them.

⁹ Although the efforts of the principal’s employees influence outcomes, we assume that the principal can compel its employees to exert their maximum effort (i.e., “your future with the firm depends on the outcome”). Thus, we focus on how to ensure the best efforts of the agent, a key factor in reaping business value. Regarding the interplay of the vendors and the consultants, vendors often form strategic alliances with major implementers (e.g., Anderson 2001) and recently have started their own consulting businesses (e.g., Pike 2006). We acknowledge the possibilities that consultants face a principal–agent problem that we do not model, but this model simplification does not represent a threat to the validity of our research design, for several reasons. First, consultants probably have a less challenging principal–agent problem because they can use specialized monitoring and management and greater internal incentives (one of the classic advantages of vendors in outsourcing; e.g., Drucker 1989). Second, vendor bears the cost of principal–agent behavior because the client holds the vendor responsible. That is, a vendor cannot use excuses, such as “we had a contract, but my staff dropped the ball so I am not responsible”. The FoxMeyer litigation attests the possibility that principals can hold agents accountable.

¹⁰ Our choice of a multiplicative rather additive model for composite risk is also in the spirit of the O-ring production functions (Kremer 1993) – failure of the space shuttle *Challenger* was caused by the malfunction of small O-rings. In our context (i.e., multiplicative rather than additive risk aggregation), EIT outcome success depends on removal of both controllable and uncontrollable risks.

el focuses purely on moral hazard issues.¹¹ At the end of each project period, the binary project 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 (binary) realization of the outcome. Conditional on a successful first period, some inherent project risks can be reduced, such as removing known obstacles to success through business process redesign. Consequently, both the principal and the vendor share a common updated belief about project risks, such that \bar{P}_2 gets updated from \bar{P}_1 and results in $\bar{P}_2 \geq \bar{P}_1$ if the first period is successful. We discuss the updating affected by project size s_i subsequently.

Figure 1 Two-Period Principal-Agent IT Contract Game



¹¹ The assumption that project value is known seems reasonable for large enterprise projects, which are similar across vendors and clients. We do not consider the issue of learning about project value (e.g., Richmond and Seidmann 1993, Whang 1992). Similarly, 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 capability 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 capability (e.g., Snir and Hitt 2004).

The possibility of learning between the two periods derives from simple 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 capability (agent expertise improves from β_1 to β_2 , with $\beta_2 \geq \beta_1$) is common knowledge. Moreover, we assume agent learning increases with project size (s_i) at a decreasing rate (concave). We also consider more elaborate formulations of agent learning as a model extension.

We employ a sub-game perfect equilibrium (SPE) concept (e.g., Lambert 1983). The principal's pure strategy is the reward function $[a_i, b_i; s_i, i = 1, 2]$, and the agent's pure strategy is mapping from the reward function to actions. The SPE strategies must satisfy the following scenario: The principal selects a set of parameters $[a_i, b_i; s_i, i = 1, 2]$ at time 0 to maximize its overall expected profit Π_i in both periods, subject to the behavior of the agent. The agent's decision variables include whether to accept the principal's contracts and, if accepted, how much effort (x_i) to exert at each period to maximize its own profit. The SPE concept restricts our attention to the class of reward functions (or contracts) in which a profit-maximizing agent participates voluntarily (which demands individual rationality or an IR constraint) and then implements the action desired by the principal (i.e., the contract must satisfy an incentive compatibility or IC constraint).

Table 1 Key Notation

Q	True value ($Q \equiv R_1 + R_2$) of the implementation project
s_i	Size of the EIT project in period i ($s_1 = \alpha, s_2 = 1 - \alpha$)
R_i	Value (revenue, cash flow) attributed to a successful IT project at period i ($R_i = s_i Q, i = 1, 2$)
a_i	Principal's fixed fee payment to the agent at period i
b_i	Scale factor of an incentive payment $b_i R_i$
$@ \equiv [a_i, b_i, s_i]$	Principal's IT contract
λ	Discount factor, same for the principal and the agent
p_i	Production function in period i $\left(p_i = \bar{P}_i (1 - e^{-\beta_i \frac{x_i}{s_i}}) \right)$

β_i	Average expertise (capability, type) in period i
x_i	Effort (cost, investment) in period i
Δ_i	One-shot in-house IT value in period i ($\Delta_i \equiv \bar{P}_i Q - 1 / \beta_i - \ln(\beta_i \bar{P}_i Q) / \beta_i$)
π_i	Agent's period i expected profit
π_i^r	Agent's period i reservation profit
π	Agent's entire project profit over both periods
π^r	Agent's entire project reservation profit over both periods
Π_i	Principal's expected profit, attributed to the IT project, for period i
V	Business value of outsourcing ($V \equiv \Pi_1 + p_1 \lambda \Pi_2$)

3.2. Solution and Discussions

We solve for the principal's optimal contracting problem in two steps. Given any project size s_i , we first solve for the optimal two-period fee structures (a_i^*, b_i^*) by backward induction, and then solve for the optimal size s_i^* , as discussed in the next section. At the beginning of Period 2, after observing the agent's first-period performance (given updated information of \bar{P}_2 and the agent's updated capability β_2), the principal maximizes expected profit in the second period by updating the contract parameters (a_2, b_2) , subject to standard agent IR and IC constraints:

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

subject to

$$(i) \pi_2(x_2^*) = p_2(a_2 + b_2 R_2 - x_2^*) + (1 - p_2)(a_2 - x_2^*) \geq \pi_2^r, \text{ and} \quad (\text{Agent IR})$$

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

The optimal solution (a_2^*, b_2^*) then becomes:

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

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

Similarly, the principal's problem at time 0 (i.e., 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) $\pi(x_1^*) = p_1(a_1 + b_1 R_1 - x_1^*) + (1 - p_1)(a_1 - x_1^*) + p_1 \lambda \pi_2^* \geq \pi^r$, and (Agent IR)
(ii) x_1^* solves $\max_{x_1} [p_1(a_1 + b_1 R_1 - x_1) + (1 - p_1)(a_1 - x_1) + p_1 \lambda \pi_2^*]$. (Agent IC)

For convenience of exposition, we define $\bar{P}_i Q - \frac{1}{\beta_i} - \frac{\ln(\beta_i \bar{P}_i Q)}{\beta_i} \equiv \Delta_i$, which equals the one-shot optimal project value provided by an in-house project team with expertise of β_i when efforts are observable. The principal's first-period optimal strategy, denoted (a_1^*, b_1^*) , becomes:

$$a_1^* = -\left[s_1 \Delta_1 + \bar{P}_1 \lambda s_2 \Delta_2 \right] + \frac{s_1}{\beta_1} \ln \left[\frac{s_1 Q + \lambda s_2 \Delta_2}{s_1 Q} \right] + \pi^r, \quad (3)$$

$$b_1^* = 1 + \frac{\lambda (s_2 \Delta_2 - \pi_2^r)}{s_1 Q} = 1 + \lambda \frac{\Pi_2^*}{R_1}. \quad (4)$$

Consequently, the principal's total IT project value is

$$V^*(@) = \Pi_1^* + p_1 \lambda \Pi_2^* = -a_1^* = \left\{ \bar{P}_1 R_1 - \frac{s_1}{\beta_1} - x_1^* \right\} + \bar{P}_1 \lambda \left[\bar{P}_2 R_2 - \frac{s_2}{\beta_2} - x_2^* \right] - \pi^r, \quad (5)$$

where $x_1^* = \frac{s_1}{\beta_1} \ln \left[\frac{\beta_1 \bar{P}_1 (s_1 Q + \lambda s_2 \Delta_2)}{s_1} \right]$ and $x_2^* = \frac{s_2}{\beta_2} \ln(\beta_2 \bar{P}_2 Q)$.

According to equation (5), $V^*(@)$ equals the expected value accruing from the system in both periods, less the agent's total profit requirement over both periods. In equilibrium and under the assumptions of risk neutrality for both firms, the optimal contract design $@ \equiv [a_i^*, b_i^*; s_i, i = 1, 2]$ is equivalent to the principal "selling" the project to the agent ($V^* = -a_1^*$, $\Pi_2^* = -a_2^*$), in support of the empirical observations by McFarlan and Nolan (1995) in real-life IT outsourcing projects. This solution also appears in other studies of IT contracting (e.g., Whang 1992). The agent nets exactly its reservation profits, because both IR constraints are binding ($\pi^* = \pi^r$, $\pi_2^* = \pi_2^r$), as is common in principal-agent models in which the principal initiates the contract offers.

These results extend repeated principal-agent results (e.g., Lambert 1983) to accommodate reward and production functions, linkages between periods, specific to EIT setting, that incorporate learning. We show in Proposition 1 that contracts parameterized with our optimal fee structure fall into the class of *viable*

contracts (Whang 1992) and achieve the Pareto-optimal outcome within a given parameter specification (equivalent to in-sourcing with an internal vendor with the same capabilities as the contractor).

PROPOSITION 1. *Given any size of an IT project s_i ($i = 1, 2$),*

(i) *There exists a unique SPE, and*

(ii) *The contract $@ \equiv [a_i^*, b_i^*; s_i]$, ($i = 1, 2$), as specified in equations (1)–(4), is viable.*

Proposition 2 characterizes some expected 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 .*

From equations (3) and (5), we know

$$V^* = -a_1^* = s_1 \Delta_1 + \bar{P}_1 \lambda s_2 \Delta_2 - \frac{s_1}{\beta_1} \ln \left[\frac{s_1 Q + \lambda s_2 \Delta_2}{s_1 Q} \right] - \pi^r. \quad (6)$$

Note immediately from equation (6) that IT project value is convex over project size in the absence of learning.¹² In addition, we compare two project structures, subdividing versus no subdividing (grouping), which is equivalent to comparing a two-stage versus a one-stage contract. We define a one-stage contract or big bang implementation as one in which there are no intermediate decision points and the project is not structured to be divisible. Although intermediate deliverables may exist, they do not provide value by themselves and do not create a contractually specified decision point. To make the comparison concise and conservative, we focus on this type of one-stage contract versus other structures.

Benchmarked with a one-stage contract, a two-staged incentive contract enables the client to extract additional effort, $\frac{s_1}{\beta_1} \ln \left[\frac{s_1 Q + \lambda s_2 \Delta_2}{s_1 Q} \right]$, from the agent in the earlier period to overcome the moral hazard

issue, which in turn creates additional value, $\bar{P}_1 \lambda s_2 (\Delta_2 - \Delta_1)$, in the later period. This relationship can be stated as in equation (7) by rearranging equation (6),

¹² Because $\frac{\partial V^2(@)}{\partial \alpha^2} = \frac{1}{\alpha \beta_i} \left[\frac{\lambda \Delta_2}{\alpha Q + \lambda (1 - \alpha) \Delta_2} \right]^2 > 0$.

$$V^* = s_1\Delta_1 + \bar{P}_1\lambda s_2\Delta_1 + \bar{P}_1\lambda s_2(\Delta_2 - \Delta_1) - \frac{s_1}{\beta_1} \ln \left[\frac{s_1Q + \lambda s_2\Delta_2}{s_1Q} \right] - \pi^r. \quad (7)$$

We summarize our model-based insights in Proposition 3 and illustrate it with several examples in online Appendix A.

PROPOSITION 3. (i) When \bar{P}_i, β_i ($i = 1, 2$) are fixed, IT value $V(@)$ is U-shaped (or convex) over project size. (ii) Subdividing (two-stage contract) is preferable to grouping (one-stage contract) if and only if \bar{P}_i, β_i ($i = 1, 2$) satisfies the following inequality: $\bar{P}_1\lambda s_2(\Delta_2 - \Delta_1) \geq \frac{s_1}{\beta_1} \ln \left[1 + \frac{\lambda s_2\Delta_2}{s_1Q} \right]$.

The assertions in Propositions 1–3 require only assumption A1; that is, two-stage contracts are preferable to one-stage contracts as long as either the revenue stream of the EIT project is large and inherent project risk decreases significantly after the successful completion of the first period, or there is a significant improvement in second-period expertise compared with initial expertise and the initial expertise of the project team is high.¹³ Without learning over time (i.e., fixed β_1, β_2), the IT value is convex with first-period project size, which yields a boundary solution in which the optimal first-period project size is as small as allowable ($\alpha^* = \underline{\alpha}$) or else an “all-out” single-period full project ($\alpha^* = 1$). Thus, a relationship among IT value, learning, and contract structure emerges. In the next section, we restrict our analysis to a two-stage contract with learning to characterize the conditions in which IT value becomes concave with first-period project size.

4. Impact of Learning Structures

In our model, we define learning as increases in expertise $\beta_2(\alpha)$ over time (or project size) α . Our second model assumption therefore states:

ASSUMPTION (A2: Concave Learning). $\beta_2(\alpha)$ is increasing and concave over time (or project size) α .

¹³ Subdividing is desirable as long as subsequent learning occurs.

We consider a learning structure in which second-period learning relates to the size of the first period, and first-period expertise β_1 is fixed, such as in the class of functions that take the form $\beta_2(\alpha) = \beta_1 + (\bar{\beta}_2 - \beta_1)\alpha^{\frac{1}{n}}$, where n is interpreted as the rate of learning ($\bar{\beta}_2$ is a constant that bounds β_2). Assumption 2 is satisfied in the case of $n \geq 1$, which accommodates both the linear and power law relationships, but is violated when $0 < n < 1$, which captures convex (rather than concave, as assumed in A2) learning over time. Assumption 2 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 apply A2 to more general cases (e.g., learning in both stages, investments in learning) with similar results.

4.1. Concave Learning Structure

Our key explanation of why concave learning (A2) may lead to a concave IT value over time relies on the direct consequence of the structure of the optimal value function. Applying the chain rule and taking total differentials of $V(@, \beta_2(\alpha))$ over α , we have:

$$\frac{dV}{d\alpha} = \frac{\partial V}{\partial \alpha} + \frac{\partial V}{\partial \beta_2(\alpha)} \frac{\partial \beta_2(\alpha)}{\partial \alpha}, \text{ and } \frac{d^2V}{d\alpha^2} = \frac{\partial^2 V}{\partial \alpha^2} + \frac{\partial^2 V}{\partial \beta_2(\alpha)^2} \left[\frac{\partial \beta_2(\alpha)}{\partial \alpha} \right]^2 + \frac{\partial V}{\partial \beta_2(\alpha)} \frac{\partial^2 \beta_2(\alpha)}{\partial \alpha^2}.$$

When β_2 is not a function of project size, the second and third term in $d^2V/d\alpha^2$ disappear, leaving only the first term, which is positive. In turn, this structure yields a convex $V(@, \beta_2(\alpha))$. When $\beta_2(\alpha)$ is time varying and concave over α , such that $\partial^2 \beta_2(\alpha) / \partial \alpha^2 < 0$, coupled with $\partial^2 V / \partial \beta_2(\alpha)^2 < 0$ and $\partial V / \partial \beta_2(\alpha) > 0$, it leads to a possible negative $d^2V/d\alpha^2$ or concave $V(@, \beta_2(\alpha))$. Theorem 1 derives precise conditions in which this state is realized.

THEOREM 1. *Given A2, IT value $V(@, \beta_2(\alpha))$ is inverted U-shaped (concave) over project size if Q is sufficiently large and β_1 is sufficiently large.*

Because Theorem 1 establishes the existence and uniqueness of optimal project sizing, any hill-climbing numerical optimization method can readily identify the solution s_i^* ($i = 1, 2$). The following corollary characterizes the impact of the rate of learning over project value and size:

COROLLARY 1. *Assume $\partial\beta_2(n)/\partial n > 0$; the faster the rate of learning, (i) the larger the IT project value¹⁴ and (ii) the smaller the optimal size of the first period.*

4.2. Time-Varying Inherent Project Risk and Risk Reduction

We previously assumed a time-invariant inherent project risk with $\bar{P}_2 \geq \bar{P}_1$, conditional on the success of the first period. However, in the first period, the inherent project risk may depend on the size, such that a larger project may be more risky, and a big bang project has the highest inherent project risk, $\bar{P}_1(\alpha = 1)$. It is equally plausible, after the successful completion of the first-period project, that the inherent project risk changes (i.e., decreases). We summarize this discussion in Assumption 3.

ASSUMPTION (A3: Inherent Project Risk and Updating). *$\bar{P}_1(\alpha)$ is decreasing and convex over time, and $\bar{P}_2(\alpha | \text{period 1 succeeds})$ is increasing and concave over time.*

A class of functions that satisfies A3 includes $\bar{P}_1(\alpha) = \frac{1-\eta}{1-\eta e^{-\alpha}}$ and $\bar{P}_2(\alpha) = \frac{1-\eta e^{-\alpha}}{1-\eta e^{-1}}$, where

$0 \leq \eta \leq 1$ is a scale factor. This specific functional form is consistent with our previous assumption of an exponential effort–outcome relationship. Note that $\bar{P}_2(\alpha | \text{period 1 succeeds}) = \bar{P}_1(1)/\bar{P}_1(\alpha)$; that is, the information updating of the second-period inherent risk $\bar{P}_2(\alpha)$ is conditional on first-period project success. Although our setting (and the more general A3) 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 for the later stages.

COROLLARY 2. *Theorem 1 holds under A3.*

¹⁴ Note that the proof of (i) in Corollary 1 does not require A2. It is also true in the case of increasing but convex agent learning, which we exploit in our experimental design.

4.3. Generalized Concave Learning Structures

Several other plausible assumptions apply to the relationships among project sizing, learning, and effort. In general, our basic conclusions are robust to these generalizations. For example, if learning occurs during the first period ($\beta_1 = \beta_1(\alpha)$), the conclusions of our preceding analyses continue to hold:

COROLLARY 3. *Theorem 1 holds when $\beta_1(\alpha)$ is increasing and concave over time.*

We therefore consider how learning might be affected not only by project size but also by effort, such as learning by doing or explicit investments in training ($t_1(\alpha)$). In addition to the cost of training, $t_1(\alpha)$ can include the coordination cost the agent incurs to collaborate with the client. This consideration leads to a more general concept of learning $\beta_2(\alpha, x_1(\alpha), t_1(\alpha))$, which requires an extension of the concave agent learning assumption:

ASSUMPTION (A4: Learning Through Training). *$t_1(\alpha)$ is increasing and concave over time. Furthermore, $\beta_2(\alpha, x_1(\alpha), t_1(\alpha))$ is increasing and concave in training $t_1(\alpha)$.*

COROLLARY 4. *Theorem 1 holds under A4.*

Corollary 4 holds 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 $@ \equiv [a_i^*, b_i^*; s_i]$ remains the same, except that we substitute π^r with $\pi^r - t_1(\alpha)$.

5. Discussion: Modeling Assumptions and Limitations

For analytical tractability and to facilitate the implementation of our model in an experimental setting, we make several specific assumptions that could affect the model predictions. Our model follows standard principal–agent formalisms and solution procedures, such as an assumption of principal and agent rationality (which we examine in the experimental section), the use of the backward induction solution method, and the standard direct revelation approach (which ensures voluntary participation by the agent and that the agent will not deviate from the optimal action). However, additional assumptions are unique to our

setting; we explore them further in this section. Specifically, we address the assumptions of principal and agent commitment, the use of a two-period game structure, and the assumption of risk neutrality.

Although our original setup assumes principal commitment, the optimal solution is robust to relaxing this assumption in the presence of learning and inherent risk reduction. Because the agent gains project-specific capabilities and the project becomes less risky after the first stage, the principal is strictly better off continuing with the agent than seeking a market alternative (without cumulative experience benefits). Moreover, the principal cannot continue a failed project, because a failed first stage means that continuation is not technically possible. Agent commitment is somewhat more critical. If the agent were not bound to continue the project, it might engage in hold-up before initiating the second stage of a multi-stage engagement by threatening to terminate the contract unless it shares in the gains from its experience improvement. In a Nash bargaining solution, this situation would require an even division of second-period surplus due to learning $\lambda s_2 (\Delta_2 - \Delta_1)$ rather than only the amount π_2' needed to satisfy the second-stage participation constraint. This new dynamic enforcement constraint $\pi_2(x_2^*) \geq 0.5s_2 (\Delta_2 - \Delta_1)$ replaces the previous agent IR in the second period, as long as $\lambda s_2 (\Delta_2 - \Delta_1) \geq 2\pi_2'$ (otherwise, the solution remains the same, due to binding of the original IR constraint). It is straightforward to show that our main results continue to hold (which is intuitive, because the agent can “raise” the previous reservation profit as high as possible, until it hits $0.5s_2 (\Delta_2 - \Delta_1)$), with the exception of second-period surplus sharing $\pi_2^* = 0.5s_2 (\Delta_2 - \Delta_1)$ and $\Pi_2^* = -a_2^* = 0.5s_2 (\Delta_2 + \Delta_1)$.

Our model focuses on two periods, though the solution structure could expand through recursion to any finite number of stages. In each stage, the principal structures the contract to extract the full surplus of that stage by selling the project to the agent, and this contract exists in each successfully reached stage.

Ultimately, it creates a mathematically complex value function, but the nature of the contract remains stable over time.¹⁵

We also have assumed that the principal and agent act as single economic actors, which separates any internal principal–agent problem from the principal agent problem between the two firms. This assumption is conventional in contract theory literature (e.g., McAfee and McMillan 1987; see also the excellent book by Bolton and Dewatripont 2005), especially in interfirm relationship applications, and appears in prior IT contracting research (Wang et al. 1997, Whang 1992). Even a substantial enterprise IT project is still small compared with the overall size of typical clients and vendors, especially if subdivided into stages.

We further assume risk neutrality by both the principal and the agent. However, our main results are robust to at least a special form of the utility function of profit, such that the agent experiences constant absolute risk aversion (CARA), as is commonly used in principal–agent modeling when the agent is risk averse.

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

In the presence of agent risk aversion, the agent’s expected utility of profit in the second period is $Eu_2(x_2) = p_2u(a_2 + b_2R_2 - x_2) + (1 - p_2)u(a_2 - x_2)$, and at the beginning of the first period (time 0), it is $Eu_0(x_1) = p_1u(b_1R_1 + a_1 - x_1) + (1 - p_1)u(a_1 - x_1) + p_1\lambda u(\pi_2^r)$.

The second-period constraints then become

$$(i) \quad E(u_2(x_2^*)) \geq u(\pi_2^r), \text{ and} \tag{Agent IR}$$

$$(ii) \quad x_2^* \text{ solves } \max_{x_2} Eu_2(x_2), \tag{Agent IC}$$

and the constraints in the principal's problem at the beginning of the first period are

¹⁵ This result contrasts sharply with the extreme case of infinitely repeated principal–agent problem, in which case the moral hazard problem is largely mitigated or overcome completely because of the infinite number of observations (Lambert 1983). For example, Radner (1985) demonstrates the effectiveness of a period review strategy using a threshold.

$$(i) \quad Eu_0(x_1^*) \geq u(\pi^r), \text{ and} \tag{Agent IR}$$

$$(ii) \quad x_1^* \text{ solves } \max_{x_1} Eu_0(x_1). \tag{Agent IC}$$

Without loss of generality and for notational and expositional simplicity,¹⁶ we let $\bar{P}_1 = \bar{P}_2 = 1$ and $\pi_i^r = 0$ ($i=1,2$) (thus $\pi^r = 0$) so that we can focus on the impact of the agent's risk tolerance (captured by γ). We denote $u(b_2R_2) \equiv w$, $u(b_1R_1) \equiv v$, and $\frac{s_i}{\beta_i} \gamma \equiv B_i$, for $i=1,2$, then add a superscript to our key parameters to denote risk attitude (A for risk-averse and N for previous risk-neutral) in the equilibrium. By applying the *certainty equivalent principle*, we solve the two-period principal-agent problem. We summarize our findings in Theorem 2. The optimal contract parameter settings for the principal are:

$$a_1^A = \frac{s_1}{\beta_1} \ln \left[\frac{1-B_1}{1+B_1} \frac{1+v}{-v} \right] + \gamma^{-1} \ln \left\{ \frac{(1+v)(1-B_1)}{\lambda(1+B_1)} \right\}, \tag{8}$$

$$b_1^A = 1 + \lambda \frac{\Pi_2^*}{R_1} + \frac{(1+v) + (1-B_1)(1+v^{-1})}{\gamma R_1}, \tag{9}$$

$$a_2^A = \gamma^{-1} \ln \frac{1+w}{B_2} - (1-B_2) \gamma^{-1} \ln \left[\frac{1-B_2}{B_2} \frac{1+w}{-w} \right], \text{ and} \tag{10}$$

$$b_2^A = 1 + \frac{(1+w) + (1-B_2)(1+w^{-1})}{\gamma R_2}. \tag{11}$$

The agent's optimal effort levels are

$$x_1^A = \frac{s_1}{\beta_1} \ln \left[\frac{1-B_1}{1+B_1} \frac{1+v}{-v} \right], \text{ and} \tag{12}$$

¹⁶ A generalization to other cases when $\bar{P}_i < 1$ ($i=1,2$) and $\pi_i^r \neq 0$ ($i=1,2$) is straightforward. Our findings do not change qualitatively, but the solutions are unnecessarily complicated. This setting enables us to highlight insights into the impact of the agent's risk attitude.

$$x_2^A = \frac{s_2}{\beta_2} \ln \left[\frac{1 - B_2}{B_2} \frac{1 + w}{-w} \right]. \quad (13)$$

THEOREM 2. Assume A5. There exists a sub-game perfect equilibrium if and only if

$$\gamma \leq \bar{\gamma} \equiv \min \left\{ \frac{\beta_1}{s_1} (1 - 2e^{-\gamma b_1 R_1}), \frac{\beta_2}{s_2} (1 - e^{-\gamma b_2 R_2}) \right\} \text{ and } \lambda \leq \bar{\lambda} \equiv \min \left\{ \frac{1}{p_1 + ((1 - p_1)e^{\gamma b_1 R_1} + p_1)(1 - p_1)} \right\}.$$

When the agent becomes less risk averse (i.e., $\gamma \rightarrow 0$), the optimal incentives converge to the risk-neutral case.

COROLLARY 5. $b_i^A \leq b_i^N$ (for $i = 1$ or 2). Furthermore, $\lim_{\gamma \rightarrow 0} b_i^A = b_i^N$.

b_i^A ($i = 1, 2$) can be computed via Taylor's series, e.g., if $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \approx 1 + x + \frac{x^2}{2}$ then

$$b_1^A \approx \frac{-(1 - B_1) + \sqrt{(1 - B_1)^2 + 2\gamma(R_1 + \lambda\Pi_2^*)(2 - B_1)}}{\gamma R_1(2 - B_1)}, \quad b_2^A \approx \frac{-(1 - B_2) + \sqrt{(1 - B_2)^2 + 2\gamma R_2(2 - B_2)}}{\gamma R_2(2 - B_2)}.$$

By making risk-neutral or specific risk-averse (CARA form) assumptions and choosing the particular, literature-motivated functional form for the production function, we characterize conditions in which our SPE exists and thereby avoid problems associated with general utility functions (Grossman and Hart 1983). We leave the situation in which both the principal and the agent are risk averse for further research. This problem is analytically intractable, but researchers could examine it using a simulation approach, and we expect further support for our results when the principal is also risk averse. The tractability of our current model is also important for proper experimental implementation. The predictions of our model appear consistent with agent behavior in our experimental settings, which provides greater confidence that we have captured the salient aspects of the problem.

6. Experimental Evidence

The results and predictions of our model build largely on various assumptions about the principal's and agent's rationality, risk attitude, and learning. To implement and validate these theoretical predictions, we test several fundamental research questions: Do consultants respond to various incentive structures? If so,

will an optimally designed contract induce the best outcome? What is the role of learning in affecting key managerial outcomes, such as the principal’s profit and the consultant’s level of effort or investment, or choices, such as the sizing of the project?

In the following experiments, subjects act as consultants playing against a computer (which represents the client or the enterprise) in a two-stage game, in which the client contracts with the consultant for an IT project.

6.1. Experimental Design and Hypotheses

We use a 3×3 design with three contract structures (one-stage fixed fee payment, two-stage suboptimal incentive payment, and two-stage optimal incentive payment) and three rates of agent learning. To make economical use of limited resources (especially subject time), we limit our attention to six of the nine treatments, which still enables us to test the key predictions of our model. All treatments have the same expected profit and parameter set (except learning rate). Table 2 summarizes our 3×3 experiments.

Table 2 Treatments

	Slow Learning ($n = 0.5$)	Medium Learning ($n = 0.9$)	Fast Learning ($n = 20$)
One-stage fixed fee payment	T_1		
Two-stage incentive payment (suboptimal)	T_2	T_4	
Two-stage incentive payment (optimal)	T_3	T_5	T_6

Our hypotheses follow directly from our model. The risk-neutral model offers optimal contract parameter settings, which are conservative estimates compared with those produced by the risk-averse model (see Section 5). In either case, agent effort should increase as we move toward the optimal incentive contract:¹⁷

H1: Stage 1 consultant effort x_1 increases from (a) T_1 to T_2 and (b) T_2 to T_3 .

H2: Stage 2 consultant effort x_2 increases from T_2 to T_3 .

¹⁷ Strictly speaking, the hypotheses should be stated in the following form: (H1a) First-period consultant effort increases from a one-stage fixed fee payment contract to a two-stage incentive payment contract. We operationalize each type of contract with a specific realization (T_1 and T_2) for experimental and expositional purposes.

Consequently, project value also should increase:

H3: *IT value $V(@)$ increases from (a) T_1 to T_2 and (b) T_2 to T_3 .*

In our second set of hypotheses, we examine the conjecture that the contract structure that is optimal for faster rates of learning induces greater consultant effort, using a learning function parameterized

by n , where $\beta_2 = \beta_1 + (\bar{\beta}_2 - \beta_1)\alpha^n$ with $\bar{\beta}_2 = 1$:

H4: *Stage 1 consultant effort x_1 increases from (a) T_3 to T_5 , (b) T_5 to T_6 , and (c) T_3 to T_6 .*

Due to increased capability, consultants optimally decrease their effort in the second stage, which yields a net cost saving:

H5: *Second-period consultant effort x_2 decreases from (a) T_3 to T_5 , (b) T_5 to T_6 , and (c) T_3 to T_6 .*

Project value is increasing with consultant capability:

H6: *IT value $V(@)$ increases from (a) T_3 to T_5 , (b) T_5 to T_6 , and (c) T_3 to T_6 .*

Finally, we compare T_4 with T_3 to test whether fast learning reduces first-period project size but retains the same profit for the client. We also examine the consultant effort level at Stage 1, which should decrease due to size reductions:

H7: *First-period consultant effort x_1 decreases from T_3 to T_4 .*

For robustness, we conducted three series of experiments with different subjects for a total of eight sessions lasting 90 minutes each. In Experiment 1, we recruited 42 subjects (7 for each treatment) from a general student pool (undergraduate students randomly recruited on campus via campus advertisements) for two sessions in which the subjects played for cash. The second experiment used 73 student subjects (13, 12, 9, 10, 16, and 13 in treatments 1–6) playing for course credits. 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 24 executives (9, 8, and 7 in treatments 2, 3, and 6, respectively) who averaged 17 years of industry experience. We summarize the parameterization of our dry and

real and run experiments in online Appendix B (Tables TB1 and TB2). We used cash payments (Experiment 1), course credits (Experiment 2), or lottery money (Experiment 3) to provide incentives to the subjects. A hypothetical currency, called Jupi, appears in all experiments. Subjects could “test drive” the system during 10 practice (or dry) runs before they participated in the 30 real runs. Although the structure of the game remains the same between dry and real runs, we used slightly different parameters.¹⁸

6.2. Findings

We summarize our findings in Tables 3 and 4. The experiments show that the structure of the IT contracts plays a fundamental role in inducing consultant effort and maximizing project value, and subjects respond to different incentive structures.

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

	Stage 1 Effort (x_1)				Stage 2 Effort (x_2)			
	Predicted	Observed			Predicted	Observed		
		Exp. 1	Exp. 2	Exp. 3		Exp. 1	Exp. 2	Exp. 3
T_1	0	0.376 (0.499)	3.323 (14.607)	N/A	N/A	N/A	N/A	N/A
T_2	4.785	6.432 (2.667)	6.872 (4.716)	6.886 (7.459)	4.915	8.266 (3.455)	8.215 (6.479)	9.413 (11.351)
T_3	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_4	4.418	5.586 (3.153)	10.937 (53.134)	N/A	11.805	12.392 (7.516)	15.109 (66.812)	N/A
T_5	8.394	12.125 (35.184)	7.183 (5.259)	N/A	6.576	14.045 (32.467)	8.244 (6.942)	N/A
T_6	8.593	14.938 (94.852)	11.008 (9.712)	10.993 (11.217)	3.099	10.969 (140.699)	9.799 (23.617)	5.835 (17.436)

The mean is shown on top of each cell; variance is shown in parenthesis.

¹⁸ Whether students can be used for theory testing is a settled debate in experimental economics, and the use of students as subjects is standard. The purpose of our series of experiments is simply to test the robustness of our theoretical findings, rather examining or comparing behaviors (students vs. executives) among experiments. By the same token, we changed incentives between Experiment 1 (money) and Experiment 2 (course credit), both of which are recommended and used widely in experimental economics literature and that satisfy the three criteria (salient, dominant, and monotonic) stated in *Induced Value Theory* (Smith 1976), which represents the guidelines for experimental economics. The incentive used for executives (Experiment 3) is money, but instead of giving everybody money (as in Experiment 1), we randomly selected a few individuals and paid them on the basis of their performance in the experiment. This approach, a lottery incentive, has been used in economics and business (Ding 2007). The benefit of a lottery incentive is that it allows researchers (with a fixed budget) to award a relatively large sum of money to a randomly selected winner. For executives, it is unlikely they will pay attention for \$10–20, as in Experiment 1, but they will likely pay attention to a reward in the range of \$200–300, even if the chance of winning is only about one in ten. Thus, the lottery incentive in Experiment 3 accommodates the nature of the subjects (executives with high opportunity cost), and we do not expect any difference due to the lottery mechanism, consistent with the literature. Finally, our design of slightly different parameterizations for the dry versus real runs follows experimental economics literature and ensures the attention of the subjects.

In Experiment 1, consultants exerted average effort levels of 6.432 in Stage 1 in the two-stage suboptimal incentive payment T_2 , compared with 0.376 in the one-stage fixed fee payment T_1 . This difference represents a significant increase ($t = 9.005, p < 0.001$). The two-stage optimal incentive payment T_3 induces the best effort of 10.257 in Stage 1, a further increase over T_2 ($t = 4.147, p < 0.001$). Meanwhile, the optimal contract T_3 induces more effort (12.775) in Stage 2, compared with 8.266 for T_2 ($t = 4.244, p < 0.001$). As a result of these greater efforts, client profits change from a net loss in T_1 to a profit in T_2 ($t = 14.331, p < 0.001$). Profits are even greater in T_3 , but these differences are not statistically significant ($t = 0.952, n.s.$).

Table 4 Summary of Predicted versus Realized Client/Consultant Profits

	Client Profit				Consultant Profit			
	Predicted	Realized			Predicted	Realized		
		Exp. 1	Exp. 2	Exp. 3		Exp. 1	Exp. 2	Exp. 3
T_1	-12.000	-10.428	3.296	N/A	12	10.273	8.677	N/A
T_2	16.880	15.676	15.609	15.745	12	11.136	9.552	9.574
T_3	22.547	17.247	16.625	16.695	12	12.352	5.650	10.459
T_4	22.547	13.952	13.476	N/A	12	12.743	4.111	N/A
T_5	25.845	18.486	18.218	N/A	12	5.486	6.609	N/A
T_6	29.020	21.466	21.335	20.652	12	-0.642	6.508	4.887

The results are similar for Experiment 2, with significant increases in effort induced as contracts come closer to the optimal structure, as well as greater profits (though no significant difference in profit occurs between T_2 and T_3). Thus, H1a, H1b, H2, and H3a are all supported. The findings for H3b move in the right direction but are not significant, which suggests that the proper design of the contract structure is more important than fine-tuning the parameters of the optimal contract.

The impact of increased consultant competency (β_2) is mostly significant for Experiment 2 but not for Experiment 1, though in both cases, the results are directionally correct. In general, the client profits from increased consultant competency, as do the consultants as a result of their competency-enabled effort savings. Subjects make only small adaptations in their Stage 2 effort when their learning rates in-

crease by a small proportion (from T_3 to T_5) but adapt significantly when moving to the high learning rate scenario (T_6). Thus, we find mixed support for H5 and H6 (especially H5a and H6a, which consider small changes in learning rates). The results are consistent with the hypotheses at statistically significant levels (generally) for Experiment 2 but only directionally consistent for Experiment 1.

Our analysis also provides partial support for H7. In Experiment 1, consultants scale back their Stage 1 effort from 10.257 (T_3) to 5.586 (T_4) in response to a reduced first-stage project size ($\alpha = 0.3$ in T_3 versus $\alpha = 0.128$ in T_4), which is significant ($t = 4.869, p < 0.001$). Similarly, in Experiment 2, consultants reduce their Stage 1 effort from 13.966 (T_3) to 10.937 (T_4), though the change is not significant ($t = 1.042, ns$). Support for this hypothesis suggests indirectly that the client can reduce IT project risk by awarding a smaller first-stage project to a fast learning team rather than the other way around.

Table 5 Summary of Results (t-statistics)

Hypotheses	Experiment 1	Experiment 2	Experiment 3
H1a: x_1 increases from T_1 to T_2	9.005***	2.882**	N/A
H1b: x_1 increases from T_2 to T_3	4.147***	3.774**	1.436*
H2: x_2 increases from T_2 to T_3	4.244***	4.070***	1.013
H3a: V increases from T_1 to T_2	14.331***	2.395*	N/A
H3b: V increases from T_2 to T_3	0.952	0.726	0.475
H4a: x_1 increases from T_3 to T_5	0.797	3.642**	N/A
H4b: x_1 increases from T_5 to T_6	0.653	3.687***	N/A
H4c: x_1 increases from T_3 to T_6	1.250	1.501*	1.636*
H5a: x_2 decreases from T_3 to T_5	0.553	4.105***	N/A
H5b: x_2 decreases from T_5 to T_6	0.619	N/A	N/A
H5c: x_2 decreases from T_3 to T_6	0.397	2.933*	2.891**
H6a: V increases from T_3 to T_5	0.941	2.010*	N/A
H6b: V increases from T_5 to T_6	2.637*	4.375***	N/A
H6c: V increases from T_3 to T_6	5.614***	6.046***	5.205***
H7: x_1 decreases from T_3 to T_4	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 * $p < 0.1$.

Industry practice and the results from Experiments 1 and 2 suggest that domain-specific contract experience could affect our experimental outcomes. To examine this argument further and demonstrate that our experiments capture behavior consistent with the actions of real technology executives, we apply the same analyses in Experiment 3 using experienced industry executives as respondents. Overall, the results from Experiment 3 are consistent with our previous findings. Although the differences are not statistically significant (perhaps due to the small sample size), this group makes greater adaptations to differential learning rates. In the learning treatments, all groups increase their investment in Stage 1, and after the Stage 2 contract had been awarded, the executives take advantage of the competency increase from T_3 to T_6 by reducing their Stage 2 effort significantly from 10.692 to 5.835 (T_3 versus T_6 , $t = 2.891$, $p < 0.01$). The client continues to benefit from this competency increase, with a profit lift from 16.695 to 20.652 ($t = 5.205$, $p < 0.001$).

In summary, our experiments reveal the following: (1) The structure of the IT contract plays a fundamental role in inducing consultant efforts and increasing IT value, because subjects respond to various incentive contract structures; (2) learning plays an important role in creating an IT value for the client and reducing consultant spending in subsequent stages; (3) subjects tend to overspend, suggesting a certain degree of risk aversion. These findings are robust across experiments, and we did not expect and did not find significant differences among subjects. An analysis of click-through data also shows no significant differences across treatments, subjects, or stages.¹⁹

7. Conclusions

We consider the optimal EIT implementation contract design in a multiperiod setting by analytically and experimentally investigating the relationship among project value, learning, risks, and project and contract structure (e.g., payment and project sizing). To the best of our knowledge, these efforts are novel and contribute significantly to IT outsourcing and contracting literature.

Theoretically, we extend previous repeated moral hazard models to incorporate (1) learning and dynamic production functions, (2) risk diversification and reduction over project periods, (3) agent risk

attitudes, and (4) linkages (contingent contract, revenue stream, and principal hold-up by the agent) between periods. These characteristics are particularly salient to large-scale EIT projects but have not been formally treated or experimentally tested in the literature. We formalize the relationship among IT value, contract choice, inherent project risks, project sizing, and learning.

In turn, our analysis yields several new insights into contract design. First, we examine the role of agent learning. In general, multiperiod contracts tend to be favorable in the presence of vendor learning or the possibility of exogenous risk reduction. Gains from subdivision and optimal project sizing depend on the rate of learning and can be large compared with the value of the project. The subdivision of projects in which agent learning occurs typically is optimal for high-value projects, even with high risks. Second, these results are robust across situations in which learning depends not just on first-period project size but also on investments in training or agent effort in the first period; our results appear to grow stronger when we consider risk-averse agents. In the absence of learning and inherent project risk reduction, single-stage full project contracts generally are preferable among risk-neutral agents. Third, we provide a theoretical basis for the conditions in which IT value takes an inverted U shape and thus provide theoretical justification for recent empirical findings about the value creation associated with EIT projects (e.g., Aral et al. 2006; Hitt et al. 2002).

Using controlled experiments, we test the theory-based relationships among IT value, learning, and project and contract structures (e.g., incentives). The model-recommended, optimal IT contract induces the best consultant effort. In addition, the proper design of the contract structure plays a more important role than fine-tuning the contract parameters, and IT value hinges on the contract structure and consultant learning. Thus, both the client and the consultant benefit from increased competency. Our experiments suggest that our theoretical results are robust to deviations from our specific assumptions about production functions, risk neutrality, or specific risk-averse (negative exponential) forms. Academically, these findings extend literature on software contracts from development (e.g., Wang et al. 1997, Whang 1992) to off-the-shelf software im-

¹⁹ Our click-through data analysis is available upon request; we omit it here for brevity.

plementation and maintenance. Practically, they have immediate implications for IT contract design and suggest significant benefits from relatively small changes to contract structures.

Although our results provide some initial insights into one aspect of contracting for enterprise software projects, tremendous opportunities remain for extending these models and empirically investigating the relationship between project performance and contract structure using laboratory or natural experiments. Our model and experiments highlight the importance of learning and its relationship with optimal project sizing, as well as the more commonly understood issues of project risk and cost structure.

References

- Anderson, E.G. 2001. Managing the impact of high market growth and learning on knowledge worker productivity and service quality. *European Journal of Operational Research* **134** 508-524.
- Aral, S., E. Brynjolfsson, D. J. Wu. 2006. Which came first, IT or productivity? The “virtuous cycle” of investment and use in extended enterprise systems. *Proceedings of the 27th International Conference on Information Systems*, Milwaukee, WI.
- Banker, R., G. Davis, S. Slaughter. 1998. Software development practices, software complexity, and software maintenance. *Management Science* **44**(4) 433-450.
- Barry, E., T. Mukhopadhyay, S. Slaughter. 2002. Software project duration and effort: An empirical study. *Information Technology and Management* **3**(1) 113-136.
- Bolton, P., M. Dewatripont. 2005. *Contract Theory*. MIT Press, Cambridge, MA.
- Brynjolfsson, E., D. Fitoussi, L. Hitt. 2006a. The information technology iceberg. Working Paper, MIT Sloan School of Management.
- Brynjolfsson, E., A. McAfee, M. Sorell, F. Zhu. 2006b. Scale without mass: Business process replication and industry dynamics. Working paper, MIT Sloan School of Management.
- Bureau of Economic Analysis. 2007. Fixed Asset Table 2.1. Available at <http://www.bea.gov/national/FA2004/SelectTable.asp>.
- Brooks, F.P. 1975. *The Mythical Man Month*. Addison Wesley, Reading, MA.
- Choudhury, V., R. Sabherwal. 2003. Portfolios of control in outsourced software development projects. *Information Systems Research* **14**(3) 291-314.
- Clemons, E., L. Hitt, E. Snir. 2000. A risk analysis framework for IT outsourcing. Wharton OPIM Working Paper.
- Cohen, M.C., N. Agrawal, V. Agrawal. 2006. Winning in the aftermarket. *Harvard Business Review* **84**(5) 129 - 138.
- Davenport, T. H. 2005. The coming commoditization of processes. *Harvard Business Review* **83**(6) 101 - 108.
- Ding, M. 2007. An incentive aligned mechanism for conjoint analysis. *Journal of Marketing Research* **44**(2) 214-223.
- DiRomualdo, A., V. Gurbaxani. 1998. Strategic intent for IT outsourcing. *Sloan Management Review* **39**(4) 67-80.
- Drucker, P. 1989. Sell the mailroom. *Wall Street Journal* (July 25).
- Gaimon, C. 1997. Planning information technology–knowledge worker systems. *Management Science* **43**(9) 1308-1328.
- Gaimon, C., K. Napoleon, G. Ozkan. 2007. Investments in workforce knowledge and technological capabilities. Working Paper, College of Management, Georgia Institute of Technology.
- Gulledge, T., and B. Khoshnevis. 1987. Production rate, learning, and program costs. A biography. *Engineering Costs and Production Economics* **11** 223-236.
- Grossman, S. J. and O. Hart. 1983. An analysis of the principal-agent problem. *Econometrica* **51**(1) 7-45.
- Hitt, L.M., D. J. Wu, X. Zhou. 2002. ERP investment: Business impact and productivity measures. *Journal of Management Information Systems* **19**(1) 71-98.

- Holmstrom, B. 1982. Moral hazard in teams. *The Bell Journal of Economics* **13** 324-340.
- Holmstrom, B., P. Milgrom. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* **55**(2) 303-328.
- James, G. 1998. Method of payment. *CIO*, October 15.
- Kirsch, L. J. 2000. Software project management: An integrated perspective for an emerging paradigm. R.W. Zmud, ed. *Framing the Domains of IT Management Research: Glimpsing the Future through the Past*, Chap. 15. Pinnaflex Educational Resources Inc., Cincinnati, Ohio, 285-304.
- Kremer, M. 1993. The O-ring theory of economic development. *Quarterly Journal of Economics* **108**(3) 551-575.
- Kulatilaka, N., L. Lin. 2006. Impact of licensing on investment and financing of technology development. *Management Science* **52**(12) 1824-1837.
- Lacity, M., R. Hirschheim. 1993. The information systems outsourcing bandwagon: Look before you leap. *Sloan Management Review* **35**(1) 72-86.
- Lacity, M., L. Willcocks. 1998. Practices in information technology outsourcing: Lessons from experience. *MIS Quarterly* **22**(3) 363-408.
- Lambert, R. 1983. Long-term contracts and moral hazard. *The Bell Journal of Economics* **14**(2) 441-452.
- Levy, F. 1965. Adaptation in the production process. *Management Science* **11**(6) B136-B154.
- Lilien, G. L., P. Kotler, K. S. Moorthy. 1992. *Marketing Models*. New Jersey: Prentice Hall.
- Loch C.H., S. Kavadias. 2002. Dynamic portfolio selection of NPD programs using marginal returns. *Management Science* **48**(10) 1227-1241.
- Markus, L. M., C. Tanis, P. C. van Fenema. 2000. Multisite ERP implementations. *Communications of the ACM* **43**(4) 42-46.
- McAfee, A. 2002. The impact of enterprise information technology adoption on operational performance: An empirical investigation. *Production and Operations Management* **11**(1) 33-53.
- McAfee, A. 2003. When too much IT knowledge is a dangerous thing. *MIT Sloan Management Review* **44**(2) 83-89.
- McAfee, A. 2006. Mastering the three worlds of information technology. *Harvard Business Review* **84**(11) 141-149.
- McAfee, R. P., J. McMillan. 1986. Bidding for contracts: A principal-agent analysis. *The RAND Journal of Economics* **17**(3) 326-338.
- McAfee, R. P., J. McMillan. 1987. Competition for agency contracts. *The RAND Journal of Economics* **18**(2) 296-307.
- McFarlan, F. W. 1981. Portfolio approach to information systems. *Harvard Business Review* **59**(5) 142-150.
- McFarlan, F. W., R. L. Nolan. 1995. How to manage an IT outsourcing alliance. *Sloan Management Review* **36**(2) 9-23.
- Mendelson, H. 2000. FoxMeyer's Delta Project. Stanford Graduate School of Business Case, Stanford University, Stanford, CA 94305.
- Nolan, R., F. W. McFarlan. 2005. Information technology and the board of directors. *Harvard Business Review* **83**(10) 96-106.
- O'Leary, D. 2002. *Enterprise Resource Planning Systems: Systems, Life Cycle, Electronic Commerce, and Risk*. New York: Cambridge University Press.

- Pike, G. 2006. Supporting business innovation while reducing technology risk. White Paper, SAP AG.
- 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, W. B., A. Seidmann. 1993. Software development outsourcing: Contract structure and business value. *Journal of Management Information Systems* **10**(1) 57-72.
- Richmond, W. B., A. Seidmann, A. B. Whinston. 1992. Incomplete contracting issues in information systems development outsourcing. *Decision Support Systems* **8**(4) 459-477.
- Rogerson, W. P. 1985. Repeated moral hazard. *Econometrica* **53**(1) 69-76.
- SAP annual report. 2005. Retrieved February 5, 2007, www.sap.com/company/investor/reports/annualreport/2005/en/lage/risiken6.html.
- Smith, V. 1976. Induced value theory. *American Economic Review* **66**(2) 274-79
- Snir, E., L. Hitt. 2004. Vendor screening in information technology contracting with a pilot project. *Journal of Organizational Computing and Electronic Commerce* **14**(1) 61-88.
- Umble, E. J., R. R. Haft, M. M. Umble. 2003. Enterprise resource planning: Implementation procedures and critical success factors. *European Journal of Operational Research* **146** 241-257.
- Wang, E. T. G., T. Barron, A. Seidmann. 1997. Contracting structures for custom software development: The impact of informational rents and uncertainty on internal development and outsourcing. *Management Science* **43**(12) 1726-1744.
- Whang, S. 1992. Contracting for software development. *Management Science* **38**(3) 307-324.

Online Appendix A: Proofs and Examples

Proof of Proposition 1

First, we slightly modify the definition of viable contracts in the literature (e.g., Whang 1992) to our context: A contract @ is viable if it satisfies the following properties:

- (1) *Efficiency*. It induces the same equilibrium decisions as the optimal in-house implementation at each period;
- (2) *Pareto Optimality: Joint Payoff Maximization*. 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 parties have positive expected payoffs at the time of contracting, so they voluntarily sign the contract.

Second, we show the satisfaction of every required property:

- (1) is satisfied because the equilibrium is subgame perfect due to the use of backward induction.
- (2) is satisfied due to $V^I = V(@) + \pi^r$, such that the combined equilibrium payoff to the contracting parties is the same as in-house implementation at time 0. It is straightforward to show that

$$V^I = s_1\Delta_1 + \bar{P}_1\lambda s_2\Delta_2 - \frac{s_1}{\beta_1} \ln \left[\frac{s_1Q + \lambda s_2\Delta_2}{s_1Q} \right]$$
 is the in-house payoff when the efforts are observable.

- (3) is satisfied because Q is assumed to be common knowledge and due to the required IC constraints;

(4) is satisfied because the IR constraints ensure that the agent gets an expected profit of $\pi^r = \pi_1^r + \lambda \pi_2^r > 0$ at the time of contracting. At the time of contracting (time 0), the principal obtains an ex ante expected IT value $V^*(@) = -a_1^* > 0$, assuming Q is sufficiently large. Thus, all ex ante profits are positive. Q.E.D.

Proof of Proposition 2

Let $y \equiv \alpha Q + \lambda(1-\alpha)\Delta_2$. From Equation (5), take first- and second-order derivatives of $V^*(@)$, and

$$\begin{aligned} \frac{\partial V}{\partial \bar{P}_1} &= \bar{P}_1 - \frac{s_1}{\beta_1} \cdot \frac{1}{y} \quad \text{and} \quad \frac{\partial^2 V}{\partial \bar{P}_1^2} = \frac{s_1}{\beta_1 \bar{P}_1^2} > 0; \\ \frac{\partial V}{\partial \bar{P}_2} &= \lambda s_2 \frac{\partial \Delta_2}{\partial \bar{P}_2} \left[\bar{P}_1 - \frac{s_1}{\beta_1} \cdot \frac{1}{y} \right] \quad \text{and} \quad \frac{\partial^2 V(@)}{\partial \bar{P}_2^2} = \lambda s_2 \frac{\partial^2 \Delta_2}{\partial \bar{P}_2^2} \left[\bar{P}_1 - \frac{s_1}{\beta_1} \cdot \frac{1}{y} \right] + \lambda s_2 \left[\frac{\partial \Delta_2}{\partial \bar{P}_2} \right]^2 \frac{s_1}{\beta_1} \cdot \frac{1}{y^3} > 0; \\ \frac{\partial V}{\partial \beta_1} &= \frac{\alpha}{\beta_1^2} - \frac{\partial x_1}{\partial \beta_1} = \frac{x_1}{\beta_1}; \quad \text{and} \\ \frac{\partial^2 V}{\partial \beta_1^2} &= \frac{s_1}{\beta_1^3} \left[-2 \ln \frac{\beta_1 \bar{P}_1 y}{s_1} + 1 \right] < 0 \text{ if } \frac{\beta_1 \bar{P}_1 y}{s_1} > \sqrt{e} \text{ (or } Q \text{ is sufficiently large)}. \end{aligned}$$

$$\begin{aligned} \frac{\partial V}{\partial \beta_2} &= \lambda s_2 \frac{\partial \Delta_2}{\partial \beta_2} \left[\bar{P}_1 - \frac{s_1}{\beta_1} \cdot \frac{1}{y} \right], \text{ and} \\ \frac{\partial^2 V}{\partial \beta_2^2} &= \lambda s_2 \left[\frac{1 - 2 \ln(\beta_2 \bar{P}_2 Q)}{\beta_2^3} \right] \left[\bar{P}_1 - \frac{s_1}{\beta_1} \cdot \frac{1}{y} \right] + \frac{s_1}{\beta_1} \left[\frac{\lambda s_2 x_2}{y \beta_2} \right]^2 < 0 \text{ if } Q \text{ is sufficiently large.} \end{aligned}$$

Then,

$$\frac{\partial V}{\partial Q} = \left[s_1 + \lambda s_2 \frac{\partial \Delta_2}{\partial Q} \right] \left[\bar{P}_1 - \frac{s_1}{\beta_1} \cdot \frac{1}{y} \right] \quad \text{and} \quad \frac{\partial V}{\partial \lambda} = s_2 \Delta_2 \left[\bar{P}_1 - \frac{s_1}{\beta_1} \cdot \frac{1}{y} \right].$$

Note that for the contract to be feasible, $x_1 > 0$, which gives $\bar{P}_1 - \frac{s_1}{\beta_1} \cdot \frac{1}{y} > 0$. Therefore,

$$\frac{\partial V}{\partial \bar{P}_1} > 0, \quad \frac{\partial V}{\partial \beta_1} > 0, \quad \frac{\partial V}{\partial \lambda} > 0.$$

Similarly, $x_2 > 0$ indicates $\beta_2 \bar{P}_2 Q > 1$, which leads to

$$\begin{aligned} \frac{\partial \Delta_2}{\partial \beta_2} &= \frac{\ln(\beta_2 \bar{P}_2 Q)}{\beta_2^2} = \frac{x_2}{\beta_2} > 0, \quad \frac{\partial \Delta_2}{\partial \bar{P}_2} = \frac{\beta_2 \bar{P}_2 Q - 1}{\beta_2 \bar{P}_2} > 0, \quad \frac{\partial \Delta_2}{\partial Q} = \frac{\beta_2 \bar{P}_2 Q - 1}{\beta_2 Q} > 0, \\ \frac{\partial^2 \Delta_2}{\partial \beta_2^2} &= \frac{1 - 2 \ln(\beta_2 \bar{P}_2 Q)}{\beta_2^3} < 0 \text{ if } Q \bar{P}_2 \beta_2 > \sqrt{e} \text{ (or } Q \text{ is sufficiently large), and } \frac{\partial^2 \Delta_2}{\partial \bar{P}_2^2} = \frac{1}{\beta_2 \bar{P}_2^2} > 0. \end{aligned}$$

Thus, $\frac{\partial V}{\partial \bar{P}_2} > 0$, $\frac{\partial V}{\partial \beta_2} > 0$, and $\frac{\partial V}{\partial Q} > 0$. Q.E.D.

Proof of Proposition 3

- (i) Holds because $\frac{\partial V^2}{\partial \alpha^2} = -\frac{\partial^2 x_1}{\partial \alpha^2} = \frac{1}{\alpha \beta_1} \left[\frac{\lambda \Delta_2}{\alpha Q + \lambda(1-\alpha)\Delta_2} \right]^2 > 0$.
- (ii) Holds because

$$\begin{aligned} & V^*(two-stage) - V^*(one-stage) \\ &= \left\{ s_1 \Delta_1 + \bar{P}_1 \lambda s_2 \Delta_2 - \frac{s_1}{\beta_1} \ln \left[\frac{s_1 Q + \lambda s_2 \Delta_2}{s_1 Q} \right] - \pi_1^r \right\} - \left\{ s_1 \Delta_1 + \bar{P}_1 \lambda s_2 \Delta_1 - \pi_1^r \right\} \\ &= \bar{P}_1 \lambda s_2 (\Delta_2 - \Delta_1) - \left[x_1^* - s_1 \frac{\ln(\beta_1 \bar{P}_1 Q)}{\beta_1} \right] = \bar{P}_1 \lambda s_2 (\Delta_2 - \Delta_1) - \frac{s_1}{\beta_1} \ln \left[\frac{s_1 Q + \lambda s_2 \Delta_2}{s_1 Q} \right]. \text{ Q.E.D.} \end{aligned}$$

Examples 1-3.

EXAMPLE 1. Assume $\beta_2 = \beta_1$ and $\bar{P}_2 = \bar{P}_1$. A one-stage contract is preferable to a two-stage contract

because the required inequality of Proposition 3 is not satisfied; $\Delta_2 = \Delta_1$, and $\ln \left[1 + \frac{\lambda s_2 \Delta_2}{s_1 Q} \right] > 0$.

EXAMPLE 2. Assume $\bar{P}_2 = \bar{P}_1 = 1$, $\lambda = 1$, and $\beta_2 > \beta_1 / (1 - \ln 2)$. Subdividing the EIT project

(by setting $s_1 = s_2 = 1/2$) is preferable to not subdividing (setting $s_1 = 1$ or $s_2 = 0$).

EXAMPLE 3. A two-stage contract is preferable to a one-stage contract (required inequality of Proposition 3 is satisfied) if (1) $\beta_2 > \beta_1$, and β_1 is sufficiently large; or (2) $\bar{P}_2 > \bar{P}_1$, and Q is sufficiently large.

Proof of Theorem 1

Denote $\Delta_2' \equiv \frac{\partial \Delta_2}{\partial \alpha}$, $\Delta_2'' \equiv \frac{\partial^2 \Delta_2}{\partial \alpha^2}$. From Equation (5), take partial derivatives of α :

$$\frac{\partial V}{\partial \alpha} = \bar{P}_1 Q + \bar{P}_1 \lambda [-\Delta_2 + (1-\alpha)\Delta_2'] - \frac{1}{\beta_1} - \frac{\partial x_1}{\partial \alpha}, \text{ and } \frac{\partial^2 V}{\partial \alpha^2} = \bar{P}_1 \lambda [-2\Delta_2' + (1-\alpha)\Delta_2''] - \frac{\partial^2 x_1}{\partial \alpha^2}.$$

Denote $y' \equiv \frac{\partial y}{\partial \alpha} = Q - \lambda \Delta_2 + \lambda(1-\alpha)\Delta_2'$ and $y'' \equiv \frac{\partial^2 y}{\partial \alpha^2} = \lambda [-2\Delta_2' + (1-\alpha)\Delta_2'']$, then

$$\frac{\partial x_1}{\partial \alpha} = \frac{x_1}{\alpha} + \frac{\alpha}{\beta_1} \frac{y'}{y} - \frac{1}{\beta_1} = \frac{x_1}{\alpha} + \frac{\alpha}{\beta_1} \left[\frac{Q - \lambda \Delta_2 + \lambda(1-\alpha)\Delta_2'}{\alpha Q + \lambda(1-\alpha)\Delta_2} \right] - \frac{1}{\beta_1}, \text{ and}$$

$$\begin{aligned}\frac{\partial^2 x_1}{\partial \alpha^2} &= \left[\frac{1}{\beta_1} \frac{y'}{y} - \frac{1}{\beta_1} \frac{1}{\alpha} \right] + \frac{1}{\beta_1} \frac{y'}{y} + \frac{\alpha}{\beta_1} \left[\frac{y'' y - y' y'}{y^2} \right] \\ &= \frac{\alpha}{\beta_1} \frac{y''}{y} - \frac{\alpha}{\beta_1} \left[\frac{y'}{y} - \frac{1}{\alpha} \right]^2 = \frac{\lambda \alpha}{\beta_1} \left[\frac{-2\Delta_2' + (1-\alpha)\Delta_2''}{\alpha Q + \lambda(1-\alpha)\Delta_2} \right] - \frac{\alpha}{\beta_1} \left[\frac{Q - \lambda\Delta_2 + \lambda(1-\alpha)\Delta_2'}{\alpha Q + \lambda(1-\alpha)\Delta_2} - \frac{1}{\alpha} \right]^2.\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{\partial^2 V}{\partial \alpha^2} &= \bar{P}_1 y'' - \frac{\partial^2 x_1}{\partial \alpha^2} = \bar{P}_1 y'' - \frac{\alpha}{\beta_1} \frac{y''}{y} + \frac{\alpha}{\beta_1} \left[\frac{y'}{y} - \frac{1}{\alpha} \right]^2 = \bar{P}_1 \lambda [-2\Delta_2' + (1-\alpha)\Delta_2''] - \frac{\partial^2 x_1}{\partial \alpha^2} \\ &= \frac{\alpha}{\beta_1} \left[\frac{Q - \lambda\Delta_2 + \lambda(1-\alpha)\Delta_2'}{\alpha Q + \lambda(1-\alpha)\Delta_2} - \frac{1}{\alpha} \right]^2 + \lambda [-2\Delta_2' + (1-\alpha)\Delta_2''] < 0 \left[\bar{P}_1 - \frac{\alpha}{\beta_1 [\alpha Q + \lambda(1-\alpha)\Delta_2]} \right].\end{aligned}\tag{A1}$$

Note that $\left. \frac{y''}{y} \right|_{Q \rightarrow \infty} = 0$ and $\left. \frac{y'}{y} \right|_{Q \rightarrow \infty} = \frac{1 - \bar{P}_2 \lambda}{\alpha + (1-\alpha)\bar{P}_2 \lambda}$, so from (A1),

$$\left. \frac{\partial^2 V(@)}{\partial \alpha^2} \right|_{Q \rightarrow \infty} = \frac{\alpha}{\beta_1} \left[\frac{1 - \bar{P}_2 \lambda}{\alpha + (1-\alpha)\bar{P}_2 \lambda} - \frac{1}{\alpha} \right]^2 + \bar{P}_1 \lambda [-2\Delta_2' + (1-\alpha)\Delta_2''] = \frac{1}{\alpha \beta_1} \left[\frac{\bar{P}_2 \lambda}{\alpha + (1-\alpha)\bar{P}_2 \lambda} \right]^2 + \bar{P}_1 y''.\tag{A2}$$

We now show that $y'' < 0$, given A2 and $Q\bar{P}_2\beta_2 > \sqrt{e}$. Notice that

$$\frac{\partial \Delta_2}{\partial \beta_2} = \frac{\ln(\beta_2 \bar{P}_2 Q)}{\beta_2^2} = \frac{x_2}{\beta_2} > 0 \text{ and } \frac{\partial^2 \Delta_2}{\partial \beta_2^2} = \frac{1 - 2 \ln(\beta_2 \bar{P}_2 Q)}{\beta_2^3} < 0.$$

Therefore,

$$\Delta_2' = \frac{\partial \Delta_2}{\partial \beta_2} \frac{\partial \beta_2}{\partial \alpha} > 0, \text{ and } \Delta_2'' = \frac{\partial^2 \Delta_2}{\partial \beta_2^2} \left[\frac{\partial \beta_2}{\partial \alpha} \right]^2 + \frac{\partial \Delta_2}{\partial \beta_2} \frac{\partial^2 \beta_2}{\partial \alpha^2} < 0. \text{ Thus, } y'' = \lambda [-2\Delta_2' + (1-\alpha)\Delta_2''] < 0.$$

If $Q \rightarrow \infty$, then $Q\bar{P}_2\beta_2 > \sqrt{e}$ is easily satisfied. If β_1 is sufficiently large, A2 becomes $\partial^2 V(@) / \partial \alpha^2 = \bar{P}_1 y'' < 0$. Q.E.D.

Proof of Corollary 1

(i) Holds because $\frac{\partial V(\alpha)}{\partial n} = \frac{\partial V(\alpha)}{\partial \beta_2} \frac{\partial \beta_2}{\partial n} > 0$. An approximation of optimal sizing, $1 - \frac{\bar{P}_1 \lambda \Delta_2 - \Delta_1}{\bar{P}_1 \lambda \Delta_2'} \equiv \tilde{\alpha}$

yields $(1 - \tilde{\alpha})\bar{P}_1 \lambda \Delta_2' \equiv [\bar{P}_1 \lambda \Delta_2 - \Delta_1]$. An increase (decrease) of n results in the increase (decrease) of the right-hand side and thus an increase (decrease) of the left-hand side to maintain the equality, which can only be achieved by a decrease (increase) of $\tilde{\alpha}$ as $\Delta_2'' < 0$. Q.E.D.

Proof of Corollary 2

The optimal contract design $@ \equiv [a_i^*, b_i^*; s_i]$ remains unchanged. Suppressing $\beta_2(\alpha)$, the IT value function can be written as $V(\alpha, \bar{P}_2(\alpha), \bar{P}_1(\alpha))$. Then,

$$\frac{dV}{d\alpha} = \frac{\partial V}{\partial \alpha} + \sum_{i=1}^2 \left[\frac{\partial V}{\partial \bar{P}_i} \frac{\partial \bar{P}_i}{\partial \alpha} \right], \text{ and } \frac{d^2 V}{d\alpha^2} = \frac{\partial^2 V}{\partial \alpha^2} + \sum_{i=1}^2 \left\{ \frac{\partial^2 V}{\partial \bar{P}_i^2} \left[\frac{\partial \bar{P}_i}{\partial \alpha} \right]^2 + \frac{\partial V}{\partial \bar{P}_i} \frac{\partial^2 \bar{P}_i}{\partial \alpha^2} \right\}.$$

In the conditions specified in Theorem 1, β_1 is sufficiently large (i.e., $\beta_1 \rightarrow \infty$), so β_2 is also sufficiently large (i.e., $\beta_2 \rightarrow \infty$), and Q is sufficiently large (i.e., $Q \rightarrow \infty$). Coupled with the first and second derivatives, as shown in the proofs of Proposition 2,

$$\left. \frac{\partial V}{\partial \bar{P}_1} \right|_{\beta_1 \rightarrow \infty} = \bar{P}_1, \quad \left. \frac{\partial^2 V}{\partial \bar{P}_1^2} \right|_{\beta_1 \rightarrow \infty} = 0, \quad \text{and} \quad \left. \frac{\partial V}{\partial \bar{P}_2} \right|_{\beta_1 \rightarrow \infty, \beta_2 \rightarrow \infty} = \lambda \bar{P}_1 s_2 Q \quad \text{and} \quad \left. \frac{\partial^2 V}{\partial \bar{P}_2^2} \right|_{\beta_1 \rightarrow \infty, \beta_2 \rightarrow \infty, Q \rightarrow \infty} = 0.$$

Thus,

$$\left. \frac{d^2 V}{d\alpha^2} \right|_{\beta_1 \rightarrow \infty, \beta_2 \rightarrow \infty, Q \rightarrow \infty} = \frac{\partial^2 V}{\partial \alpha^2} + \bar{P}_1 \frac{\partial^2 \bar{P}_1}{\partial \alpha^2} + \lambda \bar{P}_1 s_2 Q \frac{\partial^2 \bar{P}_2}{\partial \alpha^2} = \frac{\partial^2 V}{\partial \alpha^2} + \bar{P}_1 \left[\frac{\partial^2 \bar{P}_1}{\partial \alpha^2} + \lambda s_2 Q \frac{\partial^2 \bar{P}_2}{\partial \alpha^2} \right].$$

Because $\frac{\partial^2 \bar{P}_2}{\partial \alpha^2} < 0$ (see A3), when Q is sufficiently large (i.e., $Q \rightarrow \infty$), $\frac{d^2 V}{d\alpha^2} < \frac{\partial^2 V}{\partial \alpha^2} < 0$. Q.E.D.

Proof of Corollary 3

The optimal contract design $@ \equiv [a_i^*, b_i^*; s_i]$ remains unchanged, through the first-period capability $\beta_1(\alpha)$ is a function of size α . The IT value function is $V(@, \beta_2(\alpha), \beta_1(\alpha))$:

$$\frac{dV}{d\alpha} = \frac{\partial V}{\partial \alpha} + \sum_{i=1}^2 \left\{ \frac{\partial V}{\partial \beta_i(\alpha)} \frac{\partial \beta_i(\alpha)}{\partial \alpha} \right\}, \quad \text{and} \quad \frac{d^2 V}{d\alpha^2} = \frac{\partial^2 V}{\partial \alpha^2} + \sum_{i=1}^2 \left\{ \frac{\partial^2 V}{\partial \beta_i(\alpha)^2} \left[\frac{\partial \beta_i(\alpha)}{\partial \alpha} \right]^2 + \frac{\partial V}{\partial \beta_i(\alpha)} \frac{\partial^2 \beta_i(\alpha)}{\partial \alpha^2} \right\}.$$

In the conditions specified in Theorem 1, Q is sufficiently large (i.e., $Q \rightarrow \infty$), and from the proofs of Proposition 2, $\frac{\partial^2 V}{\partial \beta_i(\alpha)^2} < 0$, and $\frac{\partial V}{\partial \beta_i(\alpha)} > 0$. Coupled with $\frac{\partial^2 \beta_i(\alpha)}{\partial \alpha^2} < 0$, $i = 1, 2$, $\frac{d^2 V}{d\alpha^2} < \frac{\partial^2 V}{\partial \alpha^2} < 0$. Q.E.D.

Proof of Corollary 4

For easy of exposition, we assume β_1 is fixed. It is straightforward to generalize the results to the case of time-varying $\beta_1(\alpha)$, given Corollary 3. The optimal contract design $@ \equiv [a_i^*, b_i^*; s_i]$ remains unchanged, with the exception of modifying x_1 , as:

$$x_1 = \frac{\alpha}{\beta_1} \ln \frac{[\bar{P}_1 \beta_1(\alpha Q + \lambda(1-\alpha)\Delta_2(\beta_2(\alpha, x_1, t_1)))]}{\alpha}.$$

Notice that $\frac{\partial \beta_2}{\partial x_1} > 0$, $\frac{\partial^2 \beta_2}{\partial x_1^2} < 0$, and

$$\frac{dV}{d\alpha} = \frac{\partial V}{\partial \alpha} + \frac{\partial V}{\partial \beta_2(\alpha)} \left[\frac{\partial \beta_2}{\partial \alpha} + \frac{\partial \beta_2}{\partial x_1} \frac{\partial x_1}{\partial \alpha} + \frac{\partial \beta_2}{\partial t_1} \frac{\partial t_1}{\partial \alpha} \right] + \frac{\partial t_1}{\partial \alpha},$$

$$\begin{aligned} \frac{d^2V}{d\alpha^2} &= \frac{\partial^2V}{\partial\alpha^2} + \frac{\partial^2V}{\partial\beta_2(\alpha)^2} \left[\frac{\partial\beta_2}{\partial\alpha} + \frac{\partial\beta_2}{\partial x_1} \frac{\partial x_1}{\partial\alpha} + \frac{\partial\beta_2}{\partial t_1} \frac{\partial t_1}{\partial\alpha} \right]^2 \\ &+ \frac{\partial V}{\partial\beta_2(\alpha)} \left[\frac{\partial^2\beta_2}{\partial\alpha^2} + \frac{\partial^2\beta_2}{\partial x_1^2} \left(\frac{\partial x_1}{\partial\alpha} \right)^2 + \frac{\partial\beta_2}{\partial x_1} \frac{\partial^2 x_1}{\partial\alpha^2} + \frac{\partial^2\beta_2}{\partial t_1^2} \left(\frac{\partial t_1}{\partial\alpha} \right)^2 + \frac{\partial\beta_2}{\partial t_1} \frac{\partial^2 t_1}{\partial\alpha^2} \right] + \frac{\partial^2 t_1}{\partial\alpha^2} \end{aligned} \quad (A3)$$

In the two sufficient conditions specified in Theorem 1— β_1 is sufficiently large and Q is sufficiently large—

$$\frac{\partial x_1}{\partial\alpha} = \frac{x_1}{\alpha} + \frac{\alpha}{\beta_1} \left[\frac{Q - \lambda\Delta + \lambda(1-\alpha)\Delta'}{\alpha Q + \lambda(1-\alpha)\Delta} \right] - \frac{1}{\beta_1} = \frac{x_1}{\alpha} - \frac{1}{\beta_1} \left[\frac{\bar{P}_2\lambda}{\alpha + (1-\alpha)\bar{P}_2\lambda} \right] > 0.$$

Thus for A4, all FOCs in A3 are positive, and all SOC's in A3 are negative, so $\frac{d^2V}{d\alpha^2} < \frac{\partial^2V}{\partial\alpha^2} < 0$.

Q.E.D.

Proof of Theorem 2

Without loss of generality and for notational and expositional simplicity, we let $\bar{P}_1 = \bar{P}_2 = 1$ and $\pi_i^r = 0$ ($i=1,2$) so that we can focus on the impact of the agent's risk tolerance (captured by γ). We denote $u(b_2R_2) \equiv w$, $u(b_1R_1) \equiv v$, and $\frac{S_i}{\beta_i} \gamma \equiv B_i$, for $i=1,2$, then add a superscript to our key parameters to denote risk attitude (A for risk-averse and N for previous risk-neutral) in the equilibrium.

At the beginning of Period 2, the principal adjusts the fees (a_2, b_2) with updated information to maximize its own expected profit Π_2 :

$$\max_{a_2, b_2} \Pi_2$$

subject to

$$(i) \quad E\left(u_2(x_2^*)\right) \geq u\left(\pi_2^r\right), \text{ and} \quad (\text{Agent IR})$$

$$(ii) \quad x_2^* \text{ solves } \max_{x_2} E u_2(x_2), \quad (\text{Agent IC})$$

where the agent's expected utility of profit at Period 2 is $E u_2(x_2) = p_2 u(a_2 + b_2 R_2 - x_2) + (1 - p_2) u(a_2 - x_2)$. The principal's expected profit at the second period remains the same as in the risk-neutral case.

Applying the certainty equivalent principle, we derive the agent's risk premium π_2^0 at Period 2 as $\pi_2^0 = p_2 b_2 R_2 + \gamma^{-1} \ln[1 - p_2 - p_2 w]$. The original principal's problem is equivalent to the following:

$$\max_{a_2, b_2} \Pi_2$$

subject to

$$(i) \quad \pi_2 - \pi_2^0 = a_2 - x_2 - \gamma^{-1} \ln[1 - p_2 - p_2 w] \geq \pi_2^r, \text{ and} \quad (\text{Agent IR})$$

$$(ii) \quad x_2^* \text{ solves } \max_{x_2} (\pi_2 - \pi_2^0). \quad (\text{Agent IC})$$

$$\text{From (ii), we have } \frac{\partial p_2}{\partial x_2} = \frac{\gamma}{1+w} - p_2 \gamma.$$

Notice that $p_2 = 1 - e^{-\beta_2 \frac{x_2}{s_2}} = 1 - \frac{s_2}{\beta_2} \frac{\partial p_2}{\partial x_2}$, which gives $\frac{\partial p_2}{\partial x_2} = \frac{\beta_2}{s_2} [1 - p_2]$. Therefore,

$$p_2 = \frac{1}{1+w} \left[1 + \frac{w}{1-B_2} \right], \quad (\text{A4a})$$

$$1 - p_2 = \frac{-w}{1+w} \frac{B_2}{1-B_2}, \text{ and} \quad (\text{A4b})$$

$$x_2^A = -\frac{s_2}{\beta_2} \ln[1 - p_2] = \frac{s_2}{\beta_2} \ln \left[\frac{1-B_2}{B_2} \frac{1+w}{-w} \right].$$

The IR constraint (i) is binding, because $\partial \Pi_2 / \partial a_2 = -1 < 0$, and $\frac{\partial(\pi_2 - \pi_2^0)}{\partial a_2} > 0$. Since we normalized $\pi_2^r = 0$, this yields:

$$a_2^A = x_2^A + \gamma^{-1} \ln[1 - (1+w)p_2] = \gamma^{-1} \ln \frac{1+w}{B_2} - (1-B_2) \gamma^{-1} \ln \left[\frac{1-B_2}{B_2} \frac{1+w}{-w} \right].$$

Taking the derivatives of a_2^* w.r.t. b_2 and making use of $\frac{\partial w}{\partial b_2} = -\gamma R_2 w$, we have

$$\frac{\partial a_2^*}{\partial b_2} = R_2 \left[\frac{B_2}{1+w} - 1 \right]. \quad (\text{A5})$$

Taking derivatives of Π_2 w.r.t. b_2 and making use of equation (A5),

$$\frac{\partial \Pi_2}{\partial b_2} = -p_2 R_2 + (1-b_2) R_2 \frac{\partial p_2}{\partial b_2} - \frac{\partial a_2^*}{\partial b_2} = (1-b_2) R_2 \frac{\partial p_2}{\partial b_2} - p_2 B_2 R_2.$$

In conjunction with $\frac{\partial p_2}{\partial b_2} = \frac{B_2}{1-B_2} \frac{-\gamma R_2 w}{(1+w)^2}$ from A4a, we obtain

$$b_2^A = 1 + \frac{(1+w) + (1-B_2)(1+w^{-1})}{\gamma R_2} \quad (\text{A6})$$

Using the second order of Taylor's series, $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \approx 1 + x + \frac{x^2}{2}$, so that

$$w = -e^{-\gamma b_2 R_2} = -1 + \gamma b_2 R_2 - \frac{(\gamma b_2 R_2)^2}{2} \quad \text{and} \quad w^{-1} = -e^{\gamma b_2 R_2} = -1 - \gamma b_2 R_2 - \frac{(\gamma b_2 R_2)^2}{2},$$

then A6 becomes, $\gamma R_2 (1 - \frac{B_2}{2}) b_2^2 + [1 - B_2] b_2 - 1 = 0$.

Dropping the negative solution (which means a penalty for a successful first-period project), we have $b_2^A \approx \frac{-(1-B_2) + \sqrt{(1-B_2)^2 + 2\gamma R_2(2-B_2)}}{\gamma R_2(2-B_2)}$.

Similarly, at the beginning of Period 1 (time 0), the agent's expected utility of profit at time 0 is:

$$Eu_0(x_1) = p_1 u(b_1 R_1 + a_1 - x_1) + (1-p_1) u(a_1 - x_1) + p_1 \lambda u(\pi_2^r).$$

The principal's problem at time 0 is:

$$\max_{a_1, b_1} \Pi_1 + p_1 \lambda \Pi_2^*$$

subject to

$$(i) \quad Eu_0(x_1^*) \geq u(\pi^r), \text{ and} \tag{Agent IR}$$

$$(ii) \quad x_1^* \text{ solves } \max_{x_1} Eu_0(x_1). \tag{Agent IC}$$

Applying the certainty equivalent principle, we derive the agent's risk premium for the entire project π^0 at time 0: $\pi^0 = p_1 b_1 R_1 + p_1 \lambda \pi_2^r + \gamma^{-1} \ln \left[1 - p_1 - p_1 v - p_1 \lambda u(\pi_2^r - a_1 + x_1) \right]$. The original principal's problem is equivalent to

$$\max_{a_1, b_1} \Pi_1 + p_1 \lambda \Pi_2^*$$

subject to

$$(i) \quad \pi - \pi^0 = a_1 - x_1^* - \gamma^{-1} \ln \left[1 - p_1 - p_1 v - p_1 \lambda u(\pi_2^r - a_1 + x_1^*) \right] \geq \pi^r, \text{ and} \tag{Agent IR}$$

$$(ii) \quad x_1 \text{ solves } \max_{x_1} (\pi - \pi^0), \tag{Agent IC}$$

$$\pi - \pi^0 = a_1 - x_1 - \gamma^{-1} \ln \left[1 - p_1 - p_1 v - p_1 \lambda u(\pi_2^r - a_1 + x_1) \right]. \tag{A7}$$

Taking the FOC of equation (A7) w.r.t. x_1 , and denoting $\pi_2^r - a_1 + x_1 \equiv A$. Notice that $A = -a_1 + x_1$ since we normalized $\pi_2^r = 0$,

$$\left[1 - p_1 - p_1 v - p_1 \lambda u(A) \right] = \gamma^{-1} \left[\frac{\partial p_1}{\partial x_1} + \frac{\partial p_1}{\partial x_1} v + \frac{\partial p_1}{\partial x_1} \lambda u(A) - p_1 \gamma \lambda u(A) \right]. \tag{A8}$$

Notice that $\frac{\partial p_1}{\partial x_1} = \frac{\beta_1}{s_1} (1 - p_1)$, and equation (A8) can be rearranged as

$$-u(A) = (1 + v)(1 - B_1) + \frac{B_1 v}{1 - p_1}. \tag{A9}$$

Utilizing the following identity, $A = -\gamma^{-1} \ln \{-u(A)\}$ or $x_1 = a_1 - \gamma^{-1} \ln \{-u(A)\}$, we obtain

$$x_1^* = a_1 - \gamma^{-1} \ln \left\{ (1 + v)(1 - B_1) + \frac{B_1 v}{1 - p_1} \right\} + \gamma^{-1} \ln \lambda. \tag{A10}$$

$$\text{From equation (A10), } \frac{\partial x_1^*}{\partial p_1} = \frac{1}{-\gamma \lambda (-u(A))} \frac{B_1 v}{(1 - p_1)^2}. \tag{A11}$$

$$\text{Note as well that } \frac{\partial x_1}{\partial p_1} = \frac{s_1}{\beta_1} \frac{1}{1 - p_1}. \tag{A12}$$

$$\text{Combining equations (A11) and (A12) gives } -\lambda u(A) = \frac{-v}{1 - p_1}. \tag{A13}$$

Combining equations (A9) and (A13) yields the following optimal production function:

$$1 - p_1 = \frac{1 + B_1}{1 - B_1} \frac{-v}{1 + v} = \frac{1 + B_1}{1 - B_1} \left(\frac{1}{1 + v} - 1 \right), \tag{A14a}$$

$$p_1 = 1 + \frac{1+B_1}{1-B_1} \frac{v}{1+v} = \frac{1-B_1+2v}{(1+v)(1-B_1)} = \frac{2}{1-B_1} - \frac{1+B_1}{1-B_1} \frac{1}{1+v}. \quad (\text{A14b})$$

Note that the condition $B_1 < 1+2v$ must be satisfied, otherwise the solution will not exist because the probability of success will be negative.

The agent's optimal effort is

$$x_1^A = -\frac{s_1}{\beta_1} \ln[1-p_1^*] = -\frac{s_1}{\beta_1} \ln\left[\frac{1+B_1}{1-B_1} \frac{-v}{1+v}\right] = \frac{s_1}{\beta_1} \ln\left[\frac{1-B_1}{1+B_1} \frac{1+v}{-v}\right]. \quad (\text{A15})$$

From equations (A10) and (A15), we know

$$a_1^A = \frac{s_1}{\beta_1} \ln\left[\frac{1-B_1}{1+B_1} \frac{1+v}{-v}\right] + \gamma^{-1} \ln\left\{\frac{(1+v)(1-B_1)}{\lambda(1+B_1)}\right\}. \quad (\text{A16})$$

It can be shown that when

$$\lambda \leq \bar{\lambda} \equiv \min\left[\frac{1}{p_1 + \left[\frac{1-p_1}{-v} + p_1\right](1-p_1)}\right] < \frac{1}{p_1 + ((1-p_1) + p_1)(1-p_1)} = 1, \text{ the optimal solution from}$$

AIC will always satisfied AIR. (AIR is not binding, thus can be ignored.) Otherwise, the equilibrium does not exist. This is a key difference of the risk-averse model with the risk neutral, where agents IR constraints in both periods are binding. It can be shown that $\lim_{\gamma \rightarrow 0} \bar{\lambda} = 1$ meaning the equilibrium already exists in the risk neutral case, even with no discounting between the two periods.

The principal's optimization problem thus becomes unconstraint optimization if we substituting a_1 (equation A16) and x_1^* (equation A15) into the principal's object function, we have

$$V = \Pi_1 + p_1 \lambda \Pi_2^* = (1-b_1)p_1 R_1 - a_1 + p_1 \lambda \Pi_2^* = p_1 [R_1 + \lambda \Pi_2^*] - b_1 p_1 R_1 - a_1.$$

$$\text{The FOC gives } \frac{\partial V}{\partial b_1} = [R_1 + \lambda \Pi_2^*] \frac{\partial p_1}{\partial b_1} - p_1 R_1 - b_1 R_1 \frac{\partial p_1}{\partial b_1} - \frac{\partial a_1}{\partial b_1} = 0. \quad (\text{A17})$$

Taking the derivative of equation (A16) w.r.t. b_1 and using $\frac{\partial v}{\partial b_1} = -\gamma R_1 v$, we have

$$\frac{\partial a_1}{\partial b_1} = \frac{B_1 - v}{1+v} R_1. \quad (\text{A18})$$

We plug in equation (A18) and rewrite equation (A17) as

$$R_1 + \lambda \Pi_2^* - b_1 R_1 = \left(p_1 + \frac{B_1 - v}{1+v}\right) R_1 \frac{\partial b_1}{\partial p_1}. \quad (\text{A19})$$

Taking the derivative of equation (A14b) w.r.t. b_1 , $\frac{\partial b_1}{\partial p_1} = \frac{(1-B_1)(1+v)^2}{(1+B_1)(-\gamma R_1 v)}$.

Therefore, equation (A19) becomes the following, which implicitly determines b_1^A :

$$b_1^A = 1 + \lambda \frac{\Pi_2^*}{R_1} + \frac{(1+v) + (1-B_1)(1+v^{-1})}{\gamma R_1}. \quad (\text{A20})$$

Using Taylor's series approximation, $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \approx 1 + x + \frac{x^2}{2}$, and the approximation of b_1^A is

$$b_1^A = \frac{-(1-B_1) + \sqrt{(1-B_1)^2 + 2\gamma(R_1 + \lambda\Pi_2^*)(2-B_1)}}{\gamma R_1(2-B_1)}. \text{ Q.E.D.}$$

Proof of Corollary 5

The condition required by Theorem 2 implies $1-B_2 \geq -w$ and $1-B_1 \geq -v$.

Therefore,

$$b_2^A = 1 + \frac{(1+w) + (1-B_2)(1+w^{-1})}{\gamma R_2} \leq 1 + \frac{(1+w) - w(1+w^{-1})}{\gamma R_2} = 1 = b_2^N$$

$$b_1^A = 1 + \lambda \frac{\Pi_2^*}{R_1} + \frac{(1+v) + (1-B_1)(1+v^{-1})}{\gamma R_1} \leq 1 + \lambda \frac{\Pi_2^*}{R_1} + \frac{(1+v) - v(1+v^{-1})}{\gamma R_1} = 1 + \lambda \frac{\Pi_2^*}{R_1} = b_1^N. \text{ Q.E.D.}$$

Online Appendix B: Experiments

In this online appendix, we provide detailed design of the experiments that aimed primarily to test the assumptions and some basic predictions of our theoretical model.

Design

We introduced the game to the participants as follows:

In this experiment, you will repeatedly play a game. Each game reflects the following managerial problem. A company (client) wants to hire a consulting firm (consultant) to install a software package, but there is no guarantee the installation will be successful. However, one important thing to remember is that the more effort the consultants put into the project, the more likely the project will be successful. In the game you will play, you will be informed of the relationship between the effort spent and the probability of success.

In this experiment, you will act as the team leader for a consulting project. Your task is to determine the optimal amount of effort (in terms of money) you want to spend during each game. The more money you spend, the more likely you will be successful. On the other hand, you will make less profit as the amount of money you spend increases. As a result, it is critical and is in your best interest to determine how much effort (money) you want to spend on the project during the game.

A typical treatment (e.g., T3 in Table 2) is introduced to the participants as follows (See Figure B2 for screen shot):

In this version, the client has hired you to complete the project. The client divides the entire project into two stages (in each round). In Stage 1 (pilot), a small proportion of the project will be given to the consulting firm (you). Upon completion of Stage 1, the outcome of installation will be observed and the consultants (you) will be paid based on a

prior agreed-upon contract (for details, see instructions later). If Stage 1 is successful, the consultant (you) will be hired to complete Stage 2 (full implementation) of the project, again be paid based on a prior specified contract (may or may not be the same as in Stage 1). In each stage, you will decide how much effort (in terms of money) you want to spend to complete the project. At the end of each stage, the computer will determine the outcome of that part of the project based on the effort (money) you have spent, and you will be paid based on the outcome of the project at that stage. Please note the more effort (money) you spend on the project, the more likely the project will be successful. As a result, it is in your best interest to determine how much effort (money) you want to spend on the project during each stage. In addition to telling us the level of effort (money) you want to spend during each period, you also need to tell us how you reached that particular number.

Platform

We used the Internet (self-developed by the research team and based on MySQL and PHP) as the basic experimental platform, which enabled us to record the time spent on each task (e.g., how long, as measured by click-through data, the subject takes to decide how much to invest in each stage). After creating the experimental platform, we engaged in extensive testing and fine tuning, using feedback from an online survey (listed at the end of this online Appendix B) from a pilot test of six subjects. The state-of-the-art environments were a behavioral lab (Experiment 1) or a mobile lab (Experiments 2 and 3), which are ideal for our controlled experiments.

Subjects

We conducted three experiments for a total of eight sessions. In Experiment 1, we recruited 42 subjects (7 for each treatment) from a general student pool at a major northeast university (randomly recruited via campus advertisements) for two sessions in which the subjects played for cash. The second experiment, which comprises five sessions, used 73 student subjects (13, 12, 9, 10, 16, and 13 in treatments 1–6, respectively) playing for course credits at a major southern university. Subjects in Experiment 2 were graduate and undergraduate students who had completed half of a semester-long course on business process and/or electronic commerce and therefore were familiar with topics such as ERP systems, implementation, and outsourcing. The single session in Experiment 3 employed 24 executives (9, 8, and 7 in treatments 2, 3, and 6, respectively) who averaged 16 years of industry experience. Each session lasted 90 minutes. The first 30 minutes consisted of instructions and dry run tests (10 rounds); the remaining 60 minutes were real runs (30 rounds).

Incentives

We used cash payments (Experiment 1), course credits (Experiment 2), or lottery money (Experiment 3) to provide incentives to the subjects. We used a hypothetical currency, called Jupi, in all experiments. Subjects were shown the following instruction to provide incentives for them to play the game (see Figure B1 for screen shot):

We will use a hypothetical currency called Jupi in this experiment. The conversion rate is \$1 or 1 course point = 30 Jupi. You will be endowed with 180 Jupi at the beginning of the experiment. If you choose the optimal effort (money) to invest in a given round, the expected profit is 12 Jupi. As a result, you could earn up to 360 Jupi (from 30 rounds) in the experiment if you make all the right decisions. Plus the initial endowment, the maximum earning is 540 Jupi, which is equivalent to \$18 or 18 points in your course grade

(totaling 100 points). At the end of the experiment, your total balance in Jupi will be computed and converted to U.S. dollars or course credits and paid to you before you leave the lab.

Parameterization

Subjects had the opportunity to “test drive” the system during 10 practice runs before they participated in the 30 real runs. Although the structure of the game remains the same between dry and real runs, we used slightly different parameters, as shown in Table 2 (dry run) and 3 (real run). Subjects were shown the following instructions before the game (see Figure B1 for screen shot):

In order to familiarize you with the experiment, you will have the opportunity to complete 10 test rounds (called Dry Runs); the results from these 10 rounds do not affect your actual earning. Once you complete the test rounds, you will engage in 30 rounds of the same actual games (called Real Runs). Please note, though the structure of the games in the Dry Runs and Real Runs are identical, the parameters of the Dry Runs are different from the Real Runs.

Table B1 list the parameters used in each dry-run experiment. Real-run parameters are listed in the main text (Table 3).²⁰ The predicted IT value $V(\alpha)$ (client profit) was computed using the formula of our model.

Table B1 Parameterization for the Dry Run Experiment

	External Environment				Stage 1 Contract		Stage 2 Contract		Client Profit
	α	β_1	n	β_2	a_1	b_1	a_2	b_2	$V(\alpha)$
T1	1	0.1	0.5	1	5.000	0.000	N/A	N/A	-5
T2	0.3	0.1	0.5	0.181	4.000	0.294	2.500	0.210	4.150
T3	0.3	0.1	0.5	0.181	-11.793	1.637	-19.114	1.000	11.793
T4	0.128	0.1	0.9	0.224	-11.793	2.512	-25.413	1.000	11.793
T5	0.3	0.1	0.9	0.336	-13.765	1.785	-23.542	1.000	13.765
T6	0.3	0.1	20	0.947	-15.728	1.930	-27.911	1.000	15.728

Notes. $Q = 100, \bar{P}_1 = \bar{P}_2 = 0.5, \lambda = 1$ across all treatments. Expected profit of the consultant remains the same $\pi_i^r = 5$ ($i = 1, 2$) across all treatments. Negative a_i means the consultant is required to pay the client in exchange for the rights to implement the project. $\beta_2 = \beta_1 + (\bar{\beta}_2 - \beta_1)\alpha^n$ with $\bar{\beta}_2 = 1$.

²⁰ The consultant’s purchasing price was less than the value of the project, and the costs (x_1 and x_2) were on the same scale as the payoff. We carefully designed these parameters because if, for example, $x_1 = 0.4$ and $Q = 400$, subjects might not pay serious attention to play this game.

Table B2 Parameterization for Real Run Experiment

	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	1	0.1	0.5	1	12.000	0.000	N/A	N/A	-12
T_2	0.3	0.1	0.5	0.181	5.000	0.421	3.500	0.281	16.880
T_3	0.3	0.1	0.5	0.181	-22.547	1.894	-26.812	1.000	22.547
T_4	0.128	0.1	0.9	0.192	-22.547	3.822	-36.178	1.000	22.547
T_5	0.3	0.1	0.9	0.336	-25.845	2.061	-31.842	1.000	25.845
T_6	0.3	0.1	20	0.947	-29.020	2.222	-36.662	1.000	29.020

Notes. $Q = 100$, $\overline{P_1} = \overline{P_2} = 0.7$, and $\lambda = 1$ across all treatments. The expected consultant profit remains the same, $\pi_i^e = 12$ ($i = 1, 2$), across all treatments. Negative a_i ($i = 1, 2$) means the consultant must pay the client (or co-invest) in exchange for the rights to implement the project. β_2 is computed using $\beta_2 = \beta_1 + (\overline{\beta_2} - \beta_1)\alpha^n$, with $\overline{\beta_2} = 1$. $V(\alpha)$ is computed from Equation (5); in equilibrium, $V(@) = -a_1$. The consultant's purchasing price is smaller than the value of the project, but the costs (x_1 and x_2) are on the same scale as the payoff. These and other carefully selected parameters ensure the subjects' serious attention to the game.

Contracts

Six different payment structures correspond to our six treatments. For example, if a subject was assigned to T3, he or she would receive the following contract (see Figure B3b for screen shot); other treatments are similar but with changing parameters as listed in Table 3.

How you will be paid:

For Stage 1,

In order to earn the right to implement Stage 1, you will need to pay the client 22.5 Jupi. If the project in Stage 1 is successful, you will earn the profit of the entire project in Stage 1 (30 Jupi), plus a bonus equals to 89.4% of the profit (equal to 26.8 Jupi). This gives you a total earning of 56.8 Jupi. If the project fails, however, you will earn nothing.

Your final payoff for stage 1,

- if successful, is 56.8 - 22.5 - (money you invest in stage 1),
- and if failed, is - 22.5 - (money you invest in stage 1).

For Stage 2,

If you succeeded in Stage 1, you will proceed to Stage 2. In order to earn the right to implement Stage 2, you will need to pay the client 26.8 Jupi. If the project in Stage 2 is successful, you will earn the profit of the entire project in Stage 2 (70 Jupi). If the project fails, however, you will earn nothing.

Your final payoff for stage 2,

- if successful, is 70 - 26.8 - (money you invest in stage 2),
- and if failed, is - 26.8 - (money you invest in stage 2).

Your final payoff for this round will be the sum of your payoff in both stages.

Risks

Subjects were shown two types of risks: uncontrollable and controllable. Uncontrollable risk captures industry-wide project risk ($\bar{P}_1 = \bar{P}_2$), which is set at 0.7 in our real run experiments. The controllable risk can be reduced by the consultant's efforts (x_1, x_2), as well as by his or her competency (β_1, β_2), as captured by a decreasing returns to scale production function ($1 - e^{-\beta_i \frac{x_i}{s_i}}$), where $s_1 = \alpha$ and $s_2 = 1 - \alpha$. The overall project risk at each stage is the product of uncontrollable risk and controllable risk; namely, $p_i = \bar{P}_i (1 - e^{-\beta_i \frac{x_i}{s_i}})$. Success and failure are determined by the computer using a random mechanism that captures the chance event (see Figures B3a and B3c for screen shots).²¹

Experimental Procedure

A typical experimental session proceeded as follows:

1. Subject signs consent form if receiving cash.
2. Subject views version of the game (randomly assigned by the researchers).
3. Subject logs in.
4. Subject reads the first instruction. Experimenter further explains the instructions to make sure subjects understand.
5. Subject moves to read own version page. Experimenter further explains until the subject has no questions.
6. Subject conducts a dry run of 10 rounds of the version assigned. Figure B2 shows a screen shot of T3 (Version 3).
7. Experimenter demonstrates one example and explains until no questions remain.
8. Subject conducts the real run of 30 rounds.
9. Subject completes an online survey.
10. Subject gets paid in cash or extra course credits before leaving the lab.

Online Survey Questionnaire

1. Did you understand the instruction? If not completely, which part is confusing?
2. Do you think the task is too hard?
3. How did you determine which number you want to invest at each stage/round?
4. Did this number change over different rounds? If yes, why did you change them?
5. Do you think there was an optimal number for the investment decision? If yes, do you think you were able to identify it?
6. Any other comments you want to share with the researchers?

²¹ To determine a 0.45 probability of winning, for example, the computer randomly draws a number between 0 and 1 following an uniform distribution. If the number is smaller than 0.45, the outcome will be recorded as successful.

Figure B1 Screen shot of the experiment's Introduction. The subjects choose the specific version assigned (1–6) after reading the instructions.

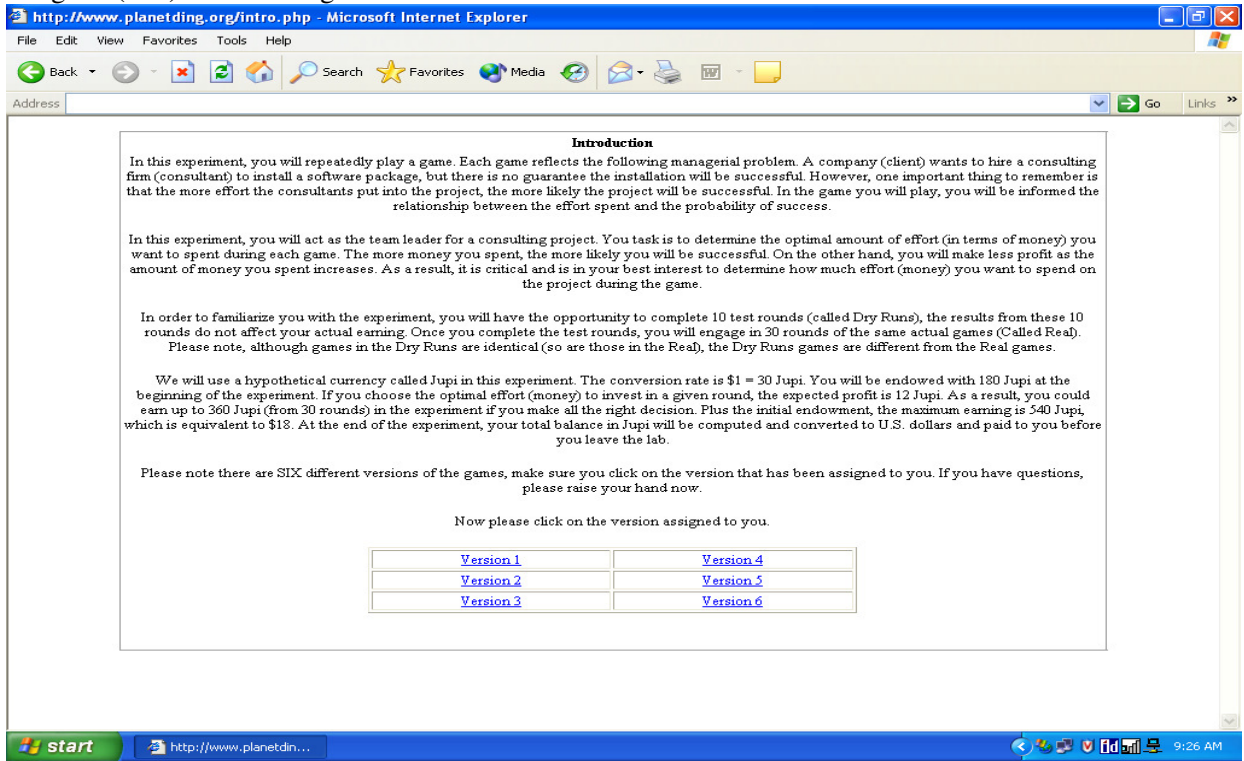


Figure B2 Screen shot of Version 3 (T3). Subjects first read specific instructions for this version, then play the dry run to familiar themselves with the game, and finally play for real.

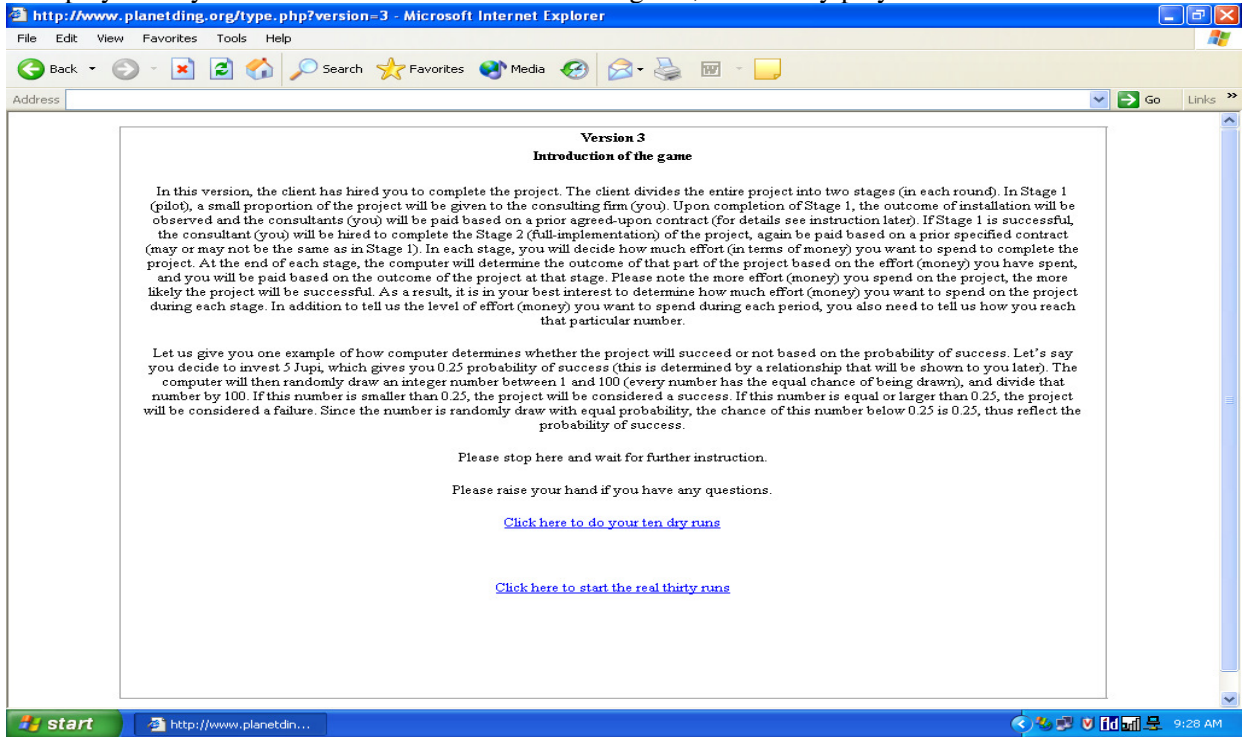


Figure B3a Screen shot of T3, Version 3, Real Run, Stage 1. Left-upper corner shows the history of the game (success/failure) to the subject; right-upper corner explains the total value of the project as well as the effort-outcome function at each stage, together with a calculated table.

History

Round	Stage 1	Result	Stage 2	Result	Jupi Made
1	10	Failure			-32.5
2	10	Failure			-32.5
3	12	Success	10	Success	55.5
4	12	Success	9	Success	56.5
5	11	Success	9	Failure	-12.5
6	11	Success	10	Success	56.5
7	11	Success	11	Failure	-14.5

**Real Run
Version 3, Stage 1**

The total profit for the project is 100 Jupi.
 The project has been divided into two Stages. In Stage 1, 30% of the project will be implemented (equals 30 Jupi) and if it is successful, the remaining 70% will be implemented in Stage 2 (equals to 70 Jupi).

The maximum probability of success if 0.7, even you spend all the effort that could possibly be spent.

The probability of success in Stage 1 (p_1), given the amount of effort (money, call it x_1) you spent in the project is

$$p_1 = 0.7 * (1 - e^{-0.1 * x_1 / 0.3})$$

The probability of success in Stage 2 (p_2), given the amount of effort (money, call it x_2) you spent in the project is

$$p_2 = 0.7 * (1 - e^{-0.181 * x_2 / 0.7})$$

To help you obtain a better understanding of this relationship, we have calculated the probabilities of success and the amount of money invested and included in the following table:

Money You Spend (x_1 or x_2 , unit Jupi)	Probability of Success in Stage 1 if this amount of money is spent in Stage 1 (p_1)	Probability of Success in Stage 2 if this amount of money is spent in Stage 2 (p_2)
0	0.0000	0.0000
1	0.1984	0.1595
2	0.3406	0.2826
3	0.4425	0.3777
4	0.5155	0.4512
5	0.5678	0.5079
6	0.6053	0.5516
7	0.6321	0.5854

Figure B3b Continued screen shot of T3. Roll down of T3, Version 3, Real Run, Stage 1. Right-bottom corner shows the contract payment terms and asks the subject for an effort for Stage 1.

16	0.6966	0.6888
17	0.6976	0.6914
18	0.6983	0.6933
19	0.6988	0.6949
20	0.6991	0.6960

How you will be paid:

For Stage 1,
 In order to earn the right to implement Stage 1, you will need to pay the client 22.5 Jupi. If the project in Stage 1 is successful, you will earn the profit of the entire project in Stage 1 (30 Jupi), plus a bonus equals 89.4% of the profit (equals to 26.8 Jupi). This gives you a total earning of 56.8 Jupi. If the project fails, however, you will earn nothing.

Your final payoff for stage 1,
 • if successful, is 56.8-22.5=(money you invest in stage 1),
 • and if failed, is -22.5=(money you invest in stage 1).

For Stage 2,
 If you succeeded in Stage 1, you will proceed to Stage 2. In order to earn the right to implement Stage 2, you will need to pay the client 26.8 Jupi. If the project in Stage 2 is successful, you will earn the profit of the entire project in Stage 2 (70 Jupi). If the project fails, however, you will earn nothing.

Your final payoff for stage 2,
 • if successful, is 70-26.8=(money you invest in stage 2),
 • and if failed, is -26.8=(money you invest in stage 2).

Your final payoff for this round will be the sum of your payoff in both stages.

Now please tell us how much money you want to spend in Stage 1:

Figure B3c Continued screen shot of T3, Version 3, Real Run, Stage 1. Outcome of Stage 1 has been realized, in this case, a success. Subject was awarded the Stage 2 contract.

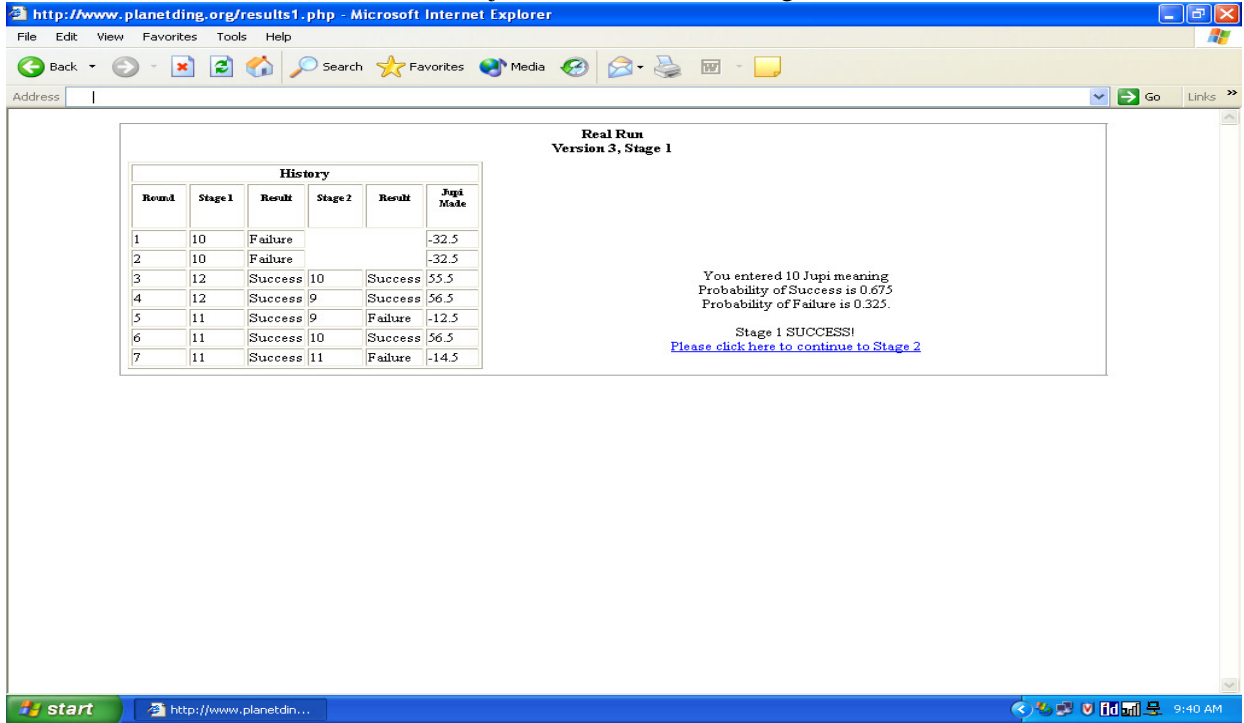


Figure B3d Continued screen shot of T3, Version 3, Real Run, Stage 2. Subject reached Stage 2 and was shown the history, as well as Stage 2 effort–outcome function (i.e., production function) and table. Subject can scroll down to put an effort for Stage 2. The computer will then inform the subject about the Stage 2 outcome, reveal the total payoff for this round, and move to the next round of the game (until all 30 runs are over).

