Uncertain Completion and Gradual Investment

Efraim Berkovich
Department of Economics
University of Pennsylvania
April 2, 2009

Abstract

Some principal-agent scenarios where there is no commitment suffer from hold-up problems—that is, projects that benefit both parties are not initiated. Research has been done on dynamic principal-agent games where gradual investment ameliorates the hold-up problem. The model here is a dynamic game where the total amount of investment to complete the project is a random variable whose realization is unknown to both principal and his agent. Because completion of the project is uncertain, the dynamics of the game evolve into a ongoing relationship. For some parameter values, the project is initiated and has positive surplus in expectation. In certain cases, however, the project is never started even though the social planner's first-best solution has positive expected surplus.

I thank Steve Matthews for reading and advising earlier versions of this paper. Robert Tayon provided helpful comments.

1 Introduction

Often, the amount of work needed to finish a project is imperfectly known. Research and development of technology and pharmaceuticals, building construction, software development, and many other undertakings of any complexity usually have only best guess estimates of the total amount of resources and time necessary for the project's successful conclusion. The total amount of work necessary may become known only once the project is actually completed, though typically the estimate of required work becomes better as more is done on the project. The decision of optimal resource allocation on these "random-sized" projects therefore can be a difficult problem to solve.

The question of optimal investment in a random-sized project becomes more complicated when a principal hires an agent to perform the work for him and the contract between the two parties is not perfectly enforceable, for example when the contract is not well-specified or when it is prohibitively expensive to prosecute violations. It is generally assumed that in the real world any moderately complex enterprise cannot be described by a complete and fully verifiable contract.

In this paper, I describe a multi-period model where a principal hires an agent to work on a random-sized project. I assume the contract is completely unenforceable and there are no future interactions between the principal and agent so that there do not exist any future payoffs to maintaining a relationship. There is a hold-up problem in that the principal has no incentive to pay the agent once the project has been successfully completed. This fact prevents the agent from exerting the efficient amount of work, and in certain circumstances may prevent the project from being initiated at all.

1.1 Motivating example

To make things a bit more concrete, consider the following scenario as a motivating example. A principal hires an agent to construct a house for him. The total amount of work to be performed is a random variable θ. Since neither the agent nor the principal can predict all the issues that will arise during construction, the realization of the random variable θ is not known to either party until after the project has been completed, but the distribution of θ is common knowledge. The principal values the completion of the house at P but loses a cost c every period (for example, insurance, mortgage payments, and other carrying costs). For simplicity, assume the principal derives no benefits unless the house is completely finished. The agent can perform an amount of work each period. The work is additive; when the cumulative amount of work through all past periods has surpassed θ, then the project has been successfully completed. Since the costs of enforcing the contract (legal fees, court time, etc.) are prohibitive, all payments by the principal are effectively discretionary.

In this scenario, construction work on the house continues so long as the probability of completion is low enough. In order to keep the probability of completion low, the agent starts to put in less work each period. The principal, mindful of this behavior by the agent, may then decide it is not worthwhile to initiate the project at all, even though both agent and principal would be better off under a social planner's solution.

1.2 Related literature

The structure of the relationship is a variation on the hold-up problem described in Pitchford and Snyder (2004) and solved through gradual discretionary payments by the principal (also Marx and Matthews (2000) for gradual investment with public goods). My model differs from Pitchford and Snyder (2004) since the uncertain completion of the project is endogenous in my model but is an exogenous construct in their paper. The structure of the described relationship requires a "self-enforcing" relational contract (Levin 2003) in that the parties maintain the relationship because of the potential for future payoffs even though the contract itself is unenforceable.

Gradual investment and Bayesian learning was described by Rob (1991) where the realization of a random variable is learned by firms entering a market and using Bayesian updating. Although Rob's model is not the hold-up variation I describe, in his model the social planner must proceed gradually and uses Bayesian updating. In my model, the social planner's solution is also a similar gradual investment strategy. This effect is in contrast to the social planner's efficient solution in many hold-up problems where the project is completed in one shot.

2 The model

There is an amount of work which is necessary to complete a project, and it is given by the random variable θ. The principal hires an agent to perform this work. Neither the principal nor the agent know the realization of θ. The game between the principal and agent runs over multiple periods where each period has three sub-periods. In the first sub-period, the agent chooses to perform some costly work. He may choose to perform zero work, in effect walking away from the job. The principal sees the amount of work performed. In the second sub-period, both parties become informed if the total amount of work has been sufficient to complete the project. Finally, in the last sub-period, the principal can pay the agent some amount b_t (which may be zero but cannot be negative). Once the project is completed, the game ends. The diagram below summarizes the timing of the stage game.

Period t

Sub-period 1	Sub-period 2	Sub-period 3
Agent performs work x_t at cost $k(x_t)$.	Both parties become informed about completion	Principal pays agent amount b_t.

The agent may end the game at the first sub-period and the principal may end the game at the third sub-period.

In each period t, the agent chooses a non-negative amount of work x_t to perform. The agent's cost function $k(x_t)$ is assumed to be continuous with $k(0) = 0$, $k' > 0$ and $k'' > 0$. Define $X_{t-1} \equiv \sum_{s=1}^{t-1} x_s$ which is the total amount of work completed before period t. The probability of the project being completed with additional work x_t depends only on X_{t-1} and is written as $\phi(X_{t-1}, x_t) \equiv \Pr(\theta \geq X_{t-1} + x_t | \theta > X_{t-1})$. The result of the agent's work is the random variable $\omega = I(X_t \geq \theta)$ which is 1 if the project completed and 0 if not. Once he sees the result of the action x_t, the principal chooses a non-negative payment b_t or ends the game unilaterally. The game has a complete information structure. The principal and agent both see each other's actions. Thus, the principal knows exactly how much work the agent performed, and the agent knows exactly how much he was paid. The distribution ϕ is common knowledge and the state of the world is also common knowledge.

Both parties are assumed to be risk-neutral, and there is no discounting. The payoff functions for the principal and the agent are derived from the following: The principal values the project's completion at P but loses a cost of $c > 0$ every period. If he terminates the project, the principal does not pay the cost c. The principal gets no benefit from a partially completed project. The agent can perform an amount of work x_t each period at a cost $k(x_t)$ and derives benefit from the payment b_t.

I consider equilibria given by time-independent strategies. Equilibrium is defined by the agent's work strategy $x(X_{t-1}, b_{t-1})$ and the principal's payment strategy $b(X_{t-1}, x_t, \omega)$; that is, strategies that depend only on variables that affect payoffs and the previous action by the opposing player.

Given the agent's strategy, the principal's problem in period t is given by the principal solving $V(X_{t-1}, x_t) = \max\left(0, \hat{V}(X_{t-1}, x_t) - c\right)$ where

$$\hat{V}(X_{t-1}, x_t) = \max_{b(\omega)} \left\{ \begin{array}{c} \phi(X_{t-1}, x_t)(P - b(1)) \\ + (1 - \phi(X_{t-1}, x_t))[V(X_t, x_{t+1}) + b(0)] \end{array} \right\}$$

where $b(\omega)$ is a discretionary payment by the principal to the agent based on whether the project completes. Since the game ends once the project is completed, the principal has no incentive to pay the agent anything, so in equilibrium $b(1) = 0$. For notational purposes, the parameter ω in the principal's strategy function can be dropped and $b(1)$ is assumed to be zero hereafter. The principal's implicit participation constraint is that $V(X_{t-1}, x_t) \geq 0$ which he can guarantee by ending the project.

Given the principal's strategy $b(X_{t-1}, x_t, \omega)$, the agent's problem at period t is given by (where for simplification, it is assumed $b(1) = 0$)

$$U(X_{t-1}) = \max_{x_t} \{(1 - \phi(X_{t-1}, x_t)) [U(X_t) + b(X_{t-1}, x_t)] - k(x_t)\}$$

The agent's implicit participation constraint is that $U(X_{t-1}) \geq 0$ which he can guarantee by doing zero work.

3 Efficient outcome

A Pareto efficient outcome is determined by maximizing the total surplus since both parties are assumed to be risk-neutral. As there are no strategy considerations, the social planner's problem is given by $\hat{W}(X_{t-1}) = \max(0, W(X_{t-1}))$ where

$$W(X_{t-1}) = \max_{x_t} \{\phi(X_{t-1}, x_t)P + (1 - \phi(X_{t-1}, x_t)) W(X_{t-1} + x_t) - c - k(x_t)\}$$

The function $W(X_{t-1})$ is the expected social surplus given total work X_{t-1} and optimal choices of x_t. Note that surplus W may be negative, in which case the project should not be continued (started).

The social planner's optimal choice of work depends on the per-period cost c. Clearly, there exists some cost $\bar{c} \leq P$, such that for all $c \geq \bar{c}$ the project will not be started at all. At the very least, a cost of $c = P$ is enough to make the project a loss. As the cost increases to the point where the social planner chooses not to start the project, the social planner does not ever choose an optimal first period work choice that decreases to zero (provided that the probability function ϕ meets certain restrictions).

Proposition 1 *If $\phi(X, x)$ is continuous and $W(X)$ exists for all X, then*
(a) $W(X)$ is continuous in X and
(b) $W(X)$ is continuous and strictly decreasing in c.

Proof. *Part (a):* To show that $W(X)$ is continuous in X, suppose that it is not. Then, for some X, $\exists\delta > 0$ such that $|W(X+\varepsilon) - W(X)| > \delta$ for some $\varepsilon \in (0, \varepsilon')$ for all $\varepsilon' > 0$. First, suppose that $W(X+\varepsilon) > W(X)$ and that x_ε^* is the optimal choice of work for $W(X+\varepsilon)$ and x^* is the optimal choice of work for $W(X)$. Then,

$$\begin{aligned} W(X+\varepsilon) &= \phi(X+\varepsilon, x_\varepsilon^*)P + (1-\phi(X+\varepsilon, x_\varepsilon^*))\,W(X+\varepsilon+x_\varepsilon^*) - k(x_\varepsilon^*) \\ &> \phi(X, x^*)P + (1-\phi(X, x^*))\,W(X+x^*) - k(x^*) + \delta = W(X) + \delta \end{aligned}$$

Now, consider $x = \varepsilon + x_\varepsilon^*$, then $\phi(X,x)P + (1-\phi(X,x))\,W(X+x) - k(x) \to \phi(X+\varepsilon, x_\varepsilon^*)P + (1-\phi(X+\varepsilon, x_\varepsilon^*))\,W(X+\varepsilon+x_\varepsilon^*) - k(x_\varepsilon^*)$ since since ϕ is continuous and k is continuous. Notice that the right hand side is just $W(X+\varepsilon)$. But then choosing $x^* = x$ gives a value to $\phi(X,x^*)P + (1-\phi(X,x^*))\,W(X+x^*) - k(x^*)$ that is at least δ higher in the limit as $\varepsilon \to 0$, thus contradicting the optimality of $W(X)$. The proof is similar for $W(X+\varepsilon) < W(X)$.

Part (b): Define $W(X; \{x_t\}, c) \equiv \phi(X, x_1)\,P + (1-\phi(X, x_1))\,W(X + x_1; \{x_t\}, c) - k(x_1) - c$ which is the welfare function of the social planner given a sequence of work choices $\{x_t\}$ and period cost c. Define by x_t^* the optimal work choices. To show that $W(X)$ decreases in c, suppose that $c_1 < c_2$ and notice that $W(X; \{x_t^{1*}\}, c_1) > W(X; \{x_t^{2*}\}, c_2)$, where $\{x_t^{i*}\}$ is the sequence of optimal work choices for the cost parameter c_i. This is so because

$$W(X; \{x_t^{2*}\}, c_2) < W(X; \{x_t^{2*}\}, c_1) \leq W(X; \{x_t^{1*}\}, c_1)$$

where the second inequality is due to $\{x_t^{1*}\}$ maximizing the valuation for cost $c_{1.}$ Since $W(X; \{x_t^{1*}\}, c_1) = W(X)$ with cost parameter c_1, W is strictly decreasing in c.

Write $W(X; \{x_t^*\}, c)$ in non-recursive form (with $\phi(X_0, x_0)$ defined to be 0 for notational purposes)

$$W(X; \{x_t^*\}, c) = \sum_{t=1}^{\infty} \left(\phi(X_t^*, x_t^*)\,P - k(x_t^*) - c\right) \prod_{s=1}^{t} \left(1 - \phi\left(X_{s-1}^*, x_{s-1}^*\right)\right)$$

Then, keeping $\{x_t^*\}$ fixed at the optimal work choices at the initial cost c, notice that $W(X; \{x_t^*\}, c)$ is linear (and so continuous) with respect to c. Let the coefficient on c be λ,

$$\lambda \equiv -\sum_{t=1}^{\infty} \prod_{s=1}^{t} \left(1 - \phi\left(X_{s-1}^*, x_{s-1}^*\right)\right)$$

To see that λ is finite, notice that $W(X; \{x_t^*\}, c) \in [0, P - c)$ so

$$\lim_{t\to\infty} \left(\phi(X_t^*, x_t^*)\,P - k(x_t^*) - c\right) \prod_{s=1}^{t} \left(1 - \phi\left(X_{s-1}^*, x_{s-1}^*\right)\right) = 0$$

Suppose that $\lim_{t\to\infty}\left(1-\phi\left(X_{t-1}^*,x_{t-1}^*\right)\right)=1$, then it must be that $\phi\left(X_t^*,x_t^*\right)P-k(x_t^*)-c\to 0$. But $\phi\left(X_t^*,x_t^*\right)\to 0$ implies that $\phi\left(X_t^*,x_t^*\right)P-k(x_t^*)-c$ becomes negative, so contradiction. Therefore, since $1-\phi\left(X_{s-1}^*,x_{s-1}^*\right)$ does not converge to one, then there exists $\varepsilon>0$ and $T<\infty$ so that

$$\prod_{s=1}^{t}\left(1-\phi\left(X_{s-1}^*,x_{s-1}^*\right)\right)<(1-\varepsilon)^t$$

for all $t>T$. So,

$$\sum_{t=1}^{\infty}\prod_{s=1}^{t}\left(1-\phi\left(X_{s-1}^*,x_{s-1}^*\right)\right)<\sum_{t=1}^{\infty}(1-\varepsilon)^t<\infty$$

Hence λ, is a negative finite number and $W(X)$ is continuous and strictly decreasing in c. ■

The following example demonstrates the social planner's solution with a history-less distribution.

Example 2 *Efficient outcome with exponential distribution*

Consider a random-sized project with a distribution of θ, such that $F(x)=1-e^{-\lambda x}$ with $\lambda>0$; this is the exponential distribution. Then for any total amount of work X and current period work x, $\phi(X,x)=\dfrac{\left(1-e^{-\lambda(X+x)}\right)-\left(1-e^{-\lambda X}\right)}{e^{-\lambda X}}=1-e^{-\lambda x}$. Note that the probability of completion is not dependent on total work done so far. The social planner's problem is $\hat{W}(X_{t-1})=\max(0,W(X_{t-1}))$ where

$$W(X_{t-1})=\max_x\left\{\left(1-e^{-\lambda x}\right)P+e^{-\lambda x}W(X_{t-1}+x)-k(x)-c\right\}$$

To maximize surplus, the social planner finds a stationary solution; that is where $x_t=x$ for all t. If W is the surplus given the optimal stationary choice of x, then

$$\begin{aligned} W &= \left(1-e^{-\lambda x}\right)P+e^{-\lambda x}W-k(x)-c \\ &\Rightarrow\ W=P-\frac{c+k(x)}{1-e^{-\lambda x}} \end{aligned}$$

The expression for W is concave because both $c+k(x)$ and $\left(1-e^{-\lambda x}\right)^{-1}$ are convex, positive, and increasing in x and therefore the negative of their product is concave. A solution can be found by the first-order condition

$$\frac{k'(x)}{c+k(x)}=\frac{\lambda e^{-\lambda x}}{1-e^{-\lambda x}}$$

The RHS is a monotonically decreasing function which goes to infinity as $x \to 0$ and goes to zero as $x \to \infty$. The LHS is finite at $x = 0$ and positive elsewhere. Therefore, in order to show that there exists a solution, it is sufficient by the intermediate value theorem to show that the LHS either does not go to zero or goes to zero more slowly than the RHS.

Showing that $\frac{k'(x)}{k(x)} > e^{-x}\left(1 - e^{-x}\right)^{-1}$ as $x \to \infty$ for all increasing convex k is sufficient for the argument. Since $k''(x) > 0$, then it must be that $k'(x)$ is bounded away from zero if $k'(x)$ was ever non-zero; that is, $k'(x) > \varepsilon$ for some $\varepsilon > 0$. Consider the limit of $e^x \frac{k'(x)}{k(x)}$ as $x \to \infty$. Since $k'(x) > \varepsilon$, then $k(x)$ goes toward infinity and $e^x k'(x)$ goes toward infinity. Therefore, using l'Hôpital's rule,

$$\begin{aligned} \lim_{x\to\infty} \frac{e^x k'(x)}{k(x)} &= \lim_{x\to\infty} \frac{\frac{d}{dx}\left(e^x k'(x)\right)}{\frac{d}{dx}\left(k(x)\right)} = \lim_{x\to\infty} \frac{e^x k'(x) + e^x k''(x)}{k'(x)} \\ &= \lim_{x\to\infty} e^x \left(1 + \frac{k''(x)}{k'(x)}\right) = \infty \end{aligned}$$

Thus, $e^x \frac{k'(x)}{k(x)} > \frac{1}{1-e^{-x}}$ for all $x > \bar{x}$ for some $\bar{x}$. Additionally, observe that even if $k''(x) = 0$ (so that k is linear), then $\frac{k'(x)}{k(x)} = o(\frac{1}{x})$ and decreases slower than $\frac{e^{-x}}{1-e^{-x}}$.

Therefore, the two curves given by the expressions on the two sides of the first-order condition equality must intersect at some point $\hat{x} < \infty$. Clearly, for a given P there exists a small enough c, such that the surplus W is positive, and in that case there exists a non-degenerate solution to the social planner's problem. The first-order condition shows that the optimal choice of work $\hat{x}$ does not depend on P. The effect on $\hat{x}$ of c and λ is ambiguous in that it depends on the shape of $k(x)$.

For a given optimal stationary choice of work $\hat{x}$, the expected number of periods necessary to complete the project is given by

$$\begin{aligned} T(\hat{x}, \lambda) &= \sum_{t=1}^{\infty} t \cdot \Pr\left(\hat{x}(t-1) < \theta \leq \hat{x}t\right) \\ &= \sum_{t=1}^{\infty} t \cdot \left(1 - e^{-\lambda \hat{x} t}\right) e^{-\lambda \hat{x}(t-1)} \\ &= \sum_{t=1}^{\infty} t \cdot \left(1 - e^{-\lambda \hat{x} t}\right) e^{-\lambda \hat{x}(t-1)} \end{aligned}$$

This infinite summation goes to 1 as $\lambda\hat{x}$ increases and goes to infinity as $\lambda\hat{x} \to 0$, though it equals zero when $\lambda\hat{x} = 0$. The summation is a discrete time version of the continuous time calculation of the mean of the exponential distribution. Note that the summation goes to $E_t[e^{-\lambda\hat{x}t}] = (\lambda\hat{x})^{-1}$ as $\lambda\hat{x} \to 0$.

4 Equilibrium

In general, the type of repeated game described here has many equilibria. The particular type of outcome I consider is where the principal's payoff is maximized, the agent's payoff is his reservation payoff of zero, and off-equilibrium path deviations are punished by "grim trigger" punishments (*i.e.*. those where the response to deviation by the opposing player is to play zero work/payment forever). The equilibrium outcomes with these restrictions are thus subgame perfect strategies. Equilibrium strategies depend only on work performed so far

Since the agent's work is observable, the principal can pay the agent just enough to meet the agent's participation constraint provided that the agent performs the optimal amount of work for the principal. If the agent does not perform the optimal amount of work, the principal can pay the agent nothing forever. If the principal fails to make a payment, the agent then refuses to work ever again. The full history of the game (b^t, x^t) is not relevant to the strategies of the players except for the total amount of work X_t and if there was a deviation in the last round. The agent's payoff in equilibrium can then be derived by solving

$$U(X_{t-1}) = \max_{x_t} \left\{ -k(x_t) + (1 - \phi(X_{t-1}, x_t)) \left(U\left(X_{t-1} + x_t\right) + b\left(X_{t-1}, x_t\right)\right) \right\}$$

where $b\left(X_{t-1}, x_t\right)$ is the principal's strategy and is given by

$$b\left(X_{t-1}, x_t\right) = \left\{ \begin{array}{c} 0 \text{ if } x_t \neq x_t^*\left(X_{t-1}\right) \\ \dfrac{k(x_t)}{1 - \phi(X_{t-1}, x_t)} \text{ if } x_t = x_t^* \end{array} \right\}$$

where $x_t^*\left(X_{t-1}\right)$ is the equilibrium choice of work. In case of deviations, one should consider that all future payments are also zero, but this fact is suppressed for notational clarity.

Proposition 3 *There exists a grim-trigger equilibrium where $U\left(0\right) > 0$ and which has the same work choices as a grim-trigger equilibrium for a lower P where $U\left(0\right) = 0$.*

Proof. This is basically a folk theorem argument. Consider an equilibrium with some P^0 where $U\left(0\right) = 0$. This equilibrium can be described by equilibrium path payments and work choices $\left(\left\{b_t^0\right\}, \left\{x_t^0\right\}\right)$. Now, consider a $P > P^0$ and consider the same work choices. The principal's expected payoff is higher. Therefore, there exists some $\delta > 0$ so that $b_{t=0} = b_0^0 + \delta$ and the principal's expected payoff is the same as before. The agent expects this payment for work x_0^0 or grim-trigger applies. The remainder of the game is the same as before. This is therefore an equilibrium. ■

Given the above result, it is in some sense sufficient to consider equilibria where the agent is held to his reservation utility of zero and the payment to the agent just covers his expected gain from working while taking into account the expected loss if the project finishes. The agent is made indifferent between working the equilibrium amount x_t^* and not working at all. Every other choice of effort leads to a utility loss because the agent gets a payment of zero but has a non-zero cost $k(x_t)$. Other equilibria where the agent has positive expected payoff are just redistributions of the total surplus and do not affect feasibility.

When the principal gets the whole surplus, the equilibrium problem can be transformed into a decision problem for the principal. After substituting $b_t = \dfrac{k(x_t)}{1-\phi(X_{t-1},x_t)}$ into his problem, the principal solves $\hat{V}(X_{t-1}) = \max\left(0, V(X_{t-1})\right)$ for the optimal equilibrium amount of work where

$$\begin{aligned} V\left(X_{t-1}\right) &= \max_{x_t}\left\{\phi(X_{t-1},x_t)P + \left(1-\phi(X_{t-1},x_t)\right)V\left(X_{t-1}+x_t\right) - k(x_t) - c\right\} \\ &\quad \text{subject to} \\ b_t &\leq V(X_{t-1}+x_t) \text{ (Incentive Constraint)} \end{aligned} \tag{1}$$

For the agent to believe the principal will make the equilibrium payment b_t^*, the principal is subject to the incentive constraint above, otherwise, the principal may choose to promise large payments in exchange for high current period work and then renege on the payment if the project did not complete that period. It is not utility maximizing for the principal to pay more than his expected utility and the agent knows this fact.

The principal's problem as defined here is the same as the social planner's problem with the addition of the incentive constraint, and so $V(X) \leq W(X)$.

Proposition 4 *If, in equilibrium (of the type described above), the principal decides to start the project, then the principal never ends the project (on the equilibrium path).*

Proof. Suppose that at some time t, the principal decides to end the project, then for any choice of labor by the agent, the continuation payoff $V\left(X_{t-1}^*+x_t\right) = 0$. Therefore, the principal has no incentive to pay the agent anything for work performed in period t. The agent knows this and therefore, he provides zero work at time t. Thus, the principal's expected payoff at time t is the per period cost c. So, the principal chooses to end the project in the previous period and get a payoff of zero instead; that is he ends the project and $V\left(X_{t-2}^{*}+x_{t-1}\right)=0$ for any choice of labor by the agent. This unraveling implies that $V(X_t^*)=0$ for all t. So, in equilibrium, if the principal decided to start the project, it must be that he never chooses to end it. ■

Notice that the above proposition is not true of the social planner's solution. Consider the case where θ is either 1 or 10 with equal probability, $k(x)=x^2$, $P=2$, and $c=\frac{1}{2}$. The first period optimal choice of work is 1. But, if the project is not finished after the first period, it is not worth it to continue, so the social planner ends the project. In the strategic outcome, this project would never be started at all. Thus, the expected surplus in the social planner's outcome is positive but is zero in the strategic outcome.

With the agent held to his reservation utility, the principal's problem is the same as the social planner's problem but with the addition of the incentive constraint. If the constraint is binding at any period t, then the principal cannot credibly induce the agent to provide the efficient amount of work, and the equilibrium amount of work must be strictly less. Because of this payment constraint, in equilibrium, the principal may choose to never start a project which has positive expected surplus in the first-best outcome.

Proposition 5 *If $\phi(X,x)$ is continuous in X and x, then there exist values for cost c such that the principal does not start the project and yet the social planner's solution provides for a strictly positive surplus.*

Proof. From results above, $W(X)$ is strictly decreasing and continuous in c. Therefore, if $W(X)=\delta>0$, then for any $\varepsilon\in(0,\delta)$, there exists some $c(\varepsilon)$ such that $W(X)=\varepsilon$. Define $W(X;x)\equiv\phi(X,x)P+(1-\phi(X,x))W(X+x)-k(x)-c$. It is continuous in x because $\phi(X,x)$ is continuous in x and because $W(X)$ is continuous in X. Notice that $W(X,x)$ is strictly decreasing in c at a fixed x because $W(X+x)$ is strictly decreasing in c and because $W(X,x)$ has a negative c term.

Note that if $W(X;x^*)>0$, then $x^*>0$. This is because $\phi(X,x)\to 0$, $k(x)\to 0$, and $W(X+x)\to W(X)$ as $x\to 0$ and so, $W(X;x)\to W(X)-c$ which is less than $W(X;x^*)=W(X)$.

Suppose $W(X;x^*)=\varepsilon$. Since $W(X,x)$ is continuous in x and since $W(X,0)=\varepsilon-c$ which is negative for sufficiently small ε, then by the intermediate value theorem, it must be that for sufficiently small ε, there exists some $\tilde{x}(\varepsilon)>0$ such that $W(X,x)<0$ for all $x<\tilde{x}(\varepsilon)$. Because the principal's optimal choice of work x^* is less than or equal to the social planner's optimal choice x^{FB} and because the principal will not continue the project if $V(X) < 0$, then the principal's optimal choice of work lies in the range $\left[\tilde{x}(\varepsilon), x^{FB}\right]$.

As c increases, $W(X,x)$ decreases continuously at every x, and so $\hat{x}(\varepsilon)$ increases continuously in c. At the same time, $W(X)$ decreases continuously in c, and so $W(X,x)$ decreases continuously as $\varepsilon = W(X)$ decreases. Therefore, $\hat{x}(\varepsilon)$ increases as ε decreases.

The incentive constraint $k(x)/(1-\phi(X,x))$ is strictly increasing in x and is strictly positive for $x > 0$. Therefore, there exists some $\varepsilon > 0$ such that the payment constraint $k(\hat{x}(\varepsilon))/(1-\phi(X,\hat{x}(\varepsilon))) > \varepsilon$ because $\tilde{x}(\varepsilon)$ (and hence the payment constraint at $\tilde{x}(\varepsilon)$) increases as ε decreases. Therefore, at this ε, $W(X) = \varepsilon$ but the payment constraint is strictly greater than ε for any value $W(X,x)$ for $x \in \left[\tilde{x}(\varepsilon), x^{FB}\right]$ since $W(X) = \varepsilon$ is the maximum of $W(X,x)$. Therefore, since $V(X) \leq W(X) = \varepsilon$, there is no x where $V(X)$ is greater than or equal to the payment constraint.

For the principal to choose to continue the project, it must be that $V(X) \geq k(x)/(1-\phi(X,x))$ for at least one x. This fact implies that there exists some c where the social planner's value function $W(X) = \varepsilon > 0$ but the principal cannot meet the payment constraint at any choice of work. ■

Example 6 *Equilibrium outcome with exponential distribution*

Consider $\phi(X,x) = 1 - e^{-\lambda x}$ as previously. Then the stationary solution to the principal's problem is found by solving

$$\begin{aligned} V &= \max_x \left\{ \left(1 - e^{-\lambda x}\right) P + e^{-\lambda x} V - k(x) - c \right\} \\ &\quad \text{subject to} \\ V &\geq k(x) e^{\lambda x} \end{aligned}$$

The first-order condition gives $\frac{k'(x)}{c+k(x)} = \frac{\lambda e^{-\lambda x}}{1-e^{-\lambda x}}$ as before. If there does not exist an x such that the payment constraint holds, then no outcome other than the "degenerate" outcome of no payment by the principal and no work by the agent is possible. For that to be the case,

$$k(x)e^{\lambda x} > P - \frac{k(x)+c}{1-e^{-\lambda x}}$$

The LHS is a strictly convex function which is non-negative and has a positive derivative at zero. The RHS is a strictly concave function and has a derivative of infinity at zero. Therefore, there is some point $\hat{x}$ where the derivative of the LHS equals the derivative of the RHS. This is the point where the two curves are the closest to each other if they do not overlap. There are always values for parameters

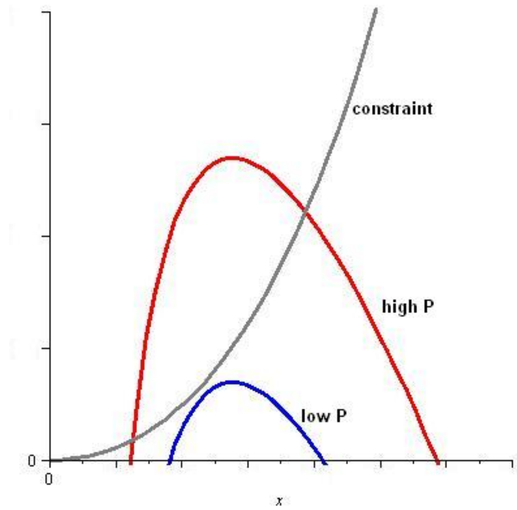

Figure 1: Hold-up for exponential example

c and P such that the expression of the RHS at $\hat{x}$ is below the expression of the LHS at $\hat{x}$ and above zero. For those parameter values, no equilibrium other than the "degenerate" one is possible, and yet the unconstrained value function is strictly positive for some range of choices of x.

The figure above shows that for the exponential example and using quadratic $k(x)$, the value function for low and high P parameters has a region above zero, so the first-best solution yields positive expected surplus. The low P value function lies below the incentive constraint and so the project is not initiated.

5 Conclusion

In this paper, I present a multi-period game model where a principal hires an agent to perform work on a project. One feature of the model is that the total amount of work needed to complete the project is random and unknown to either party. In the first-best outcome, the amount of work to be performed each period may be a complex calculation that takes into account the probability of finishing the project. Since increased work is increasing costly to the agent, there is a benefit of reducing work each period. However, as each period the project is not complete is costly to the principal, there is a benefit to trying to do more work to complete the project quickly.

Strategic play of the game is further complicated by the existence of an incentive constraint—the principal cannot credibly offer to pay the agent more than the principal's expected payoff. For certain parameters, strategic play has no equilibrium other than the degenerate one of no work while the efficient solution provides for a strictly positive surplus. Gradual payments may, therefore, be insufficient as a means of resolving the hold-up.

6 References

1. Levin, J. "Relational incentive contracts." *American Economic Review,* v.93 (2003) pp.835-847.

2. Marx, L. and S. Matthews. "Dynamic voluntary contribution to a public project." *Review of Economic Studies,* v.67 (2000) pp.327-358

3. Pitchford, R. and C. Snyder. "A solution to the hold-up problem involving gradual investment." *Journal of Economic Theory,* v. 114 (2004) pp. 88-103.

4. Rob, R. "Learning and Capacity Expansion under Demand Uncertainty." *Review of Economic Studies,* v. 58 (1991) pp. 655-675