DISCUSSION PAPER SERIES

No. 6725

REPEATED MORAL HAZARD, LIMITED LIABILITY, AND RENEGOTIATION

Susanne Ohlendorf and Patrick W. Schmitz

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

REPEATED MORAL HAZARD, LIMITED LIABILITY, AND RENEGOTIATION

Susanne Ohlendorf, University of Bonn
Patrick W. Schmitz, University of Cologne and CEPR

Discussion Paper No. 6725
February 2008

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 6725

February 2008

ABSTRACT

Repeated Moral Hazard, Limited Liability, and Renegotiation*

We consider a repeated moral hazard problem, where both the principal and the wealth-constrained agent are risk-neutral. In each of two periods, the principal can make an investment and the agent can exert unobservable effort, leading to success or failure. Incentives in the second period act as carrot and stick for the first period, so that effort is higher after a success than after a failure. If renegotiation cannot be prevented, the principal may prefer a project with lower returns; i.e., a project may be "too good" to be financed or, similarly, an agent can be "overqualified."

JEL Classification: C73 and D86
Keywords: dynamic moral hazard, hidden actions and limited liability

Susanne Ohlendorf
University of Bonn
Wirtschaftspolitische Abteilung and BGSE
Adenauerallee 24-42
53113 Bonn
GERMANY
Email: susanne.ohlendorf@uni-bonn.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=167823

Patrick W. Schmitz
University of Cologne
Department of Economics
Albertus-Magnus-Platz
50923 Köln
GERMANY
Email: patrick.schmitz@uni-koeln.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=149900

* We have benefited from very helpful discussions with Patrick Bolton, Mathias Dewatripont, Tymofiy Mylovanov, Andreas Roider, and Urs Schweizer. Financial support by Deutsche Forschungsgemeinschaft, SFB/TR15, is gratefully acknowledged.

Submitted 19 February 2008

1 Introduction

This paper offers a new perspective on repeated moral hazard problems. Consider a risk-neutral principal, who can invest to install a project and hire a risk-neutral but wealth-constrained agent. The agent can exert unobservable effort, which increases the likelihood of success. In the one-shot problem, there is a well-known trade-off between effort incentives and rent extraction, which leads to a downward distorted effort level compared to the first-best solution. We extend the standard model by assuming that there is a second period, in which the principal can again make an investment to continue the project and the agent can again exert unobservable effort. It turns out that there are several interesting insights that so far have escaped the literature on repeated moral hazard, which was focused on the case of risk-averse agents.

In particular, if the principal can commit not to renegotiate, the second period incentives can be used to partially circumvent the limited liability constraint. In the second period, the principal implements a particularly high effort level following a first-period success and a particularly low effort level following a first-period failure. The prospect of a higher second-period rent following a first-period success motivates the agent to exert more effort in the first period; i.e., rents in the second period act as reward and punishment for the first period. It should be emphasized that we assume no impact of a first-period success or failure on the second-period technology. Nevertheless, if an outsider observed today a principal-agent pair that was successful and another identical pair that was not successful, he would be right to predict that the first pair also is more likely to succeed tomorrow. In other words, a "hot hand" effect is generated endogenously, merely based on incentive considerations.[1]

Just as in the one-shot model, effort levels are distorted and not every project

[1] The term "having a hot hand" originated in basketball and means having a streak of successes that cannot be attributed to normal variation in performance. It seems to spectators that the probability of a success increases after a row of successes, even though the trials in question are independent; see Gilovich, Vallone, and Tversky (1985).

that would be installed in a first-best world will be pursued under moral hazard. It also is still the case that the principal will always prefer a project (or, equivalently, an agent) that can yield a larger return (among otherwise identical projects or agents). Somewhat surprisingly, however, the latter observation is no longer true if renegotiation cannot be ruled out.

The "hot hand" effect implies that a principal would sometimes like to commit to terminate a project following a first-period failure, even though technologically the success probability of the second period is not affected by the first-period outcome. Yet, the threat to terminate may not be credible if renegotiation cannot be prevented. In this case, a new kind of inefficiency occurs, that to the best of our knowledge has not been identified in the repeated moral hazard literature so far: The principal might deliberately choose a project that is commonly known to yield smaller potential returns than another (otherwise identical) project that is also available. Similarly, she might deliberately hire an agent that is commonly known to be less qualified.

The reason that a project might be "too good" to be funded or an agent might be "overqualified" is the fact that the principal cannot resist the temptation to renegotiate if the potential return is too attractive, which is anticipated by the agent, whose incentives to work hard in the first period are dulled. In contrast, a less qualified agent or an agent working on a less attractive project may well be willing to exert more effort, because he knows that in case of a failure he will not get a second chance. Since the credible threat to terminate the project after a first-period failure improves first-period incentives, there are indeed parameter constellations under which a relatively bad project is funded, while a better project is not.

The literature on repeated moral hazard problems and renegotiation has different strands. Most papers consider repeated versions of the traditional moral hazard setting, where the agent is risk-averse and there is a trade-off between insurance and incentives.[2] In a pioneering paper, Rogerson (1985) considered a

[2] For comprehensive surveys, see Chiappori et al. (1994) and Bolton and Dewatripont (2005,

two-period moral hazard problem and showed that the optimal second-period incentives depend on the first-period outcome, even though the periods are technologically independent. However, his result is driven by the consumption-smoothing motive of the risk-averse agent,[3] which is absent in our setting. In moral hazard models with a risk-averse agent, renegotiation is an issue even in the one-shot problem, because after the agent has chosen an effort level, there is no need to expose him to further risk. Fudenberg and Tirole (1990), Ma (1991, 1994) and Matthews (1995, 2001) show that it depends on the details of the renegotiation game whether or not effort incentives are reduced.[4] In contrast, in our framework there is scope for renegotiation only if the moral hazard problem is repeated, and the details of the renegotiation game are irrelevant for our results.

Although we consider a repeated moral hazard problem, it is interesting to note that our results are also related to the repeated adverse selection literature.[5] Specifically, in a seminal paper Dewatripont and Maskin (1995) consider a two-period model where the agent has private information about the quality of a project that he submits for funding. Ex ante, the principal would like to terminate bad projects after the first period in order to deter the agent from submitting them ("hard budget constraint"). Yet, at the beginning of the second period she is tempted to refinance them ("soft budget constraint"). The absence of commitment power thus enables bad projects to be funded. However, as has

chapter 10). Note that while focusing on different questions, the case of a risk-neutral agent has also been considered by Hirao (1993), Wang (2000), and Tamada and Tsai (2007).

[3]Cf. Malcomson and Spinnewyn (1988), Fudenberg, Holmström, and Milgrom (1990), and Rey and Salanié (1990). See also Mukoyama and Şahin (2005) for a model in which the second-period outcome is directly influenced by the first-period effort.

[4]See also Hermalin and Katz (1991) and Dewatripont et al. (2003), who consider observable but unverifiable effort, and Gigler and Hemmer (2007), who allow for an infinite number of rounds of renegotiation.

[5]The fact that the one-shot moral hazard model with a risk-neutral but wealth-constrained agent has some similarities to the one-shot adverse selection model has already been noted by Laffont and Martimort (2002, p. 147).

been pointed out by Kornai, Maskin, and Roland (2003, p. 1110), the principal would not finance a bad project if she knew the quality ex ante. In contrast, in our model a bad project may be funded, while a better project may not be funded, even though the quality is common knowledge.

The remainder of the paper is organized as follows. In Section 2.1, we introduce the one-shot moral hazard problem with a risk-neutral but wealth-constrained agent, which now is sometimes called "efficiency wage" model.[6] This model serves as a benchmark for the dynamic analysis. We then introduce the two-period model in Section 2.2.[7] In section 3, we analyze the commitment scenario and highlight the "hot hand" effect. In section 4, it is assumed that renegotiation cannot be ruled out, which leads to the "overqualification" effect. Finally, concluding remarks follow in Section 5. All proofs have been relegated to the appendix.

2 The model

2.1 The one-shot contracting problem

As a useful benchmark, let us first take a brief look at the one-shot moral-hazard problem that will be repeated twice in our full-fledged model. There are two parties, a principal and an agent, both of whom are risk-neutral. The agent has

[6] See Tirole (1999, p. 745) or Laffont and Martimort (2002, p. 174). See also Innes (1990), Pitchford (1998), and Tirole (2001) for more detailed discussions of the one-shot moral hazard model with risk-neutrality and resource constraints. Moreover, cf. the traditional efficiency wage literature (Shapiro and Stiglitz, 1984) and the literature on deferred compensation (Lazear, 1981), which are related but have a focus different from the contract-theoretic literature to which we add in the present contribution.

[7] Dynamic models with risk-neutral agents, hidden actions, and wealth constraints include also Crémer (1995), Baliga and Sjöström (1998), Che and Yoo (2001), and Schmitz (2005). Yet, they rely on features (private information about productivity, observable yet unverifiable effort, common shocks, and technological relations between the periods, respectively) which are absent in the repeated (pure) moral hazard problem studied here.

no resources of his own, so that all payments to the agent have to be nonnegative. The parties' reservation utilities are given by zero. At some initial date 0, the principal can decide whether or not to pursue a project. If she installs the project, she incurs costs I and she offers a contract to the agent. Having accepted the contract, the agent exerts unobservable effort $e \in [0,1]$ at date 1. His disutility from exerting effort is given by $c(e)$. Finally, at date 2, either a success ($y = 1$) or a failure ($y = 0$) is realized, where $\Pr\{y = 1|e\} = e$. The principal's verifiable return is given by yR.

Assumption 1. The effort cost function satisfies[8]

a) $c \in \mathcal{C}^3([0,1])$ and $c'(1) > R$,

b) $c' > 0$, $c'' > 0$, $c''' \geq 0$,

c) $c(0) = 0$ and $c'(0) = 0$.

The first-best effort level maximizes the expected total surplus $S(e) := eR - c(e)$ and is characterized by

$$S'(e^{FB}) = R - c'(e^{FB}) = 0.$$

In a first-best world, the project would be installed whenever $S(e^{FB}) \geq I$.

The principal could attain the first-best outcome, but in order to do so she would have to leave all of her returns to the agent, because payments to the agent must not be negative. Hence, the principal faces a trade-off between increasing the pie and getting a larger share for herself. To find the second-best solution, observe first that the principal will not pay anything when no revenue is generated. Next, let t denote the principal's transfer payment to the agent in case of success. The agent's expected payoff from exerting effort e is $et - c(e)$. If $t \leq R$, which obviously is optimal, the agent's maximization problem has

[8] The assumptions allow us to focus on first-order conditions. With minor notational modifications of the proofs, we could relax Assumption 1a) by including $c \in \mathcal{C}^3([0,1))$, where $\lim_{e\to 1} c'(e) > R$.

an interior solution characterized by $t = c'(e)$. Hence, the principal maximizes $P(e) := e[R - c'(e)]$ over e. The first-order condition characterizing the effort level she will implement thus is

$$P'(e^{SB}) = R - c'(e^{SB}) - e^{SB}c''(e^{SB}) = 0.$$

Note that the function P is concave, positive for $e < e^{FB}$ and negative for $e > e^{FB}$. We also define $A(e) := ec'(e) - c(e)$, the agent's rent from a contract that leads him to choose effort e. Since its derivative is $A'(e) = ec''(e)$, this is a non-negative and strictly increasing function. Hence, a higher implemented effort level yields higher rents for the agent. In order to reduce the agent's rent, the principal thus introduces a downward distortion of the induced effort level, $e^{SB} < e^{FB}$.

In the one-shot problem, the principal will install the project whenever $P(e^{SB}) \geq I$; i.e., not all projects that would be pursued in a first-best world will actually be installed. However, given the choice between two (otherwise identical) projects with possible returns $R = R_g$ and $R = R_b < R_g$, the principal will never prefer the bad project that can yield R_b only.

2.2 The two-period model

Now we turn to the full-fledged two-period model. For simplicity, we neglect discounting. At date 0, the principal decides whether $(x_1 = 1)$ or not $(x_1 = 0)$ to invest an amount $I_1 > 0$ in order to install the project and run it for the first period. If she pursues the project, she makes a take-it-or-leave-it contract offer to the agent. Having accepted the offer, at date 1 the agent chooses an unobservable first-period effort level $e_1 \in [0, 1]$, incurring disutility $c(e_1)$. At date 2, the verifiable first-period return $y_1 R$ is realized, where $y_1 \in \{0, 1\}$ denotes failure or success, and $\Pr\{y_1 = 1|e_1\} = e_1$. The project may then be terminated $(x_2(y_1) = 0)$ or continued $(x_2(y_1) = 1)$, which is verifiable.[9] In order to continue

[9] We assume that it is too costly for the principal to replace the agent at date 2, because at that point in time the parties are "locked-in" (i.e., the relationship has undergone Williamson's

the project, the principal must invest an amount $I_2 \leq I_1$.[10] In this case, at date 3 the agent chooses an unobservable second-period effort level $e_2(y_1) \in [0,1]$. Finally, at date 4 the verifiable second-period return $y_2 R$ is realized, where $y_2 \in \{0,1\}$ and $\Pr\{y_2 = 1|e_2(y_1)\} = e_2(y_1)$, and the contractually specified transfer payments are made. Note that the two periods are independent; in particular, we do not assume any technological spillovers that would make a second-period success more likely after a first-period success.

The sequence of events is illustrated in Figure 1.

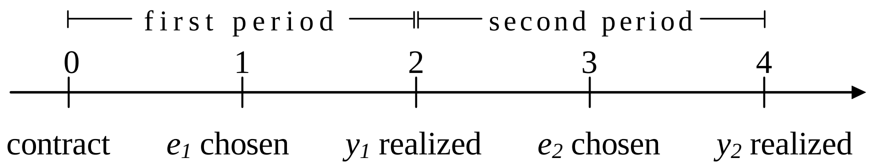

Figure 1. The sequence of events.

The first-best solution. Assume for a moment that effort were verifiable. If $S(e^{FB}) + S(e^{FB}) \geq I_1 + I_2$, then the principal would install the project ($x_1 = 1$), she would continue regardless of the first-period outcome ($x_2(0) = x_2(1) = 1$), and she would implement $e_1 = e_2(0) = e_2(1) = e^{FB}$ with a straightforward forcing contract, leaving no rent to the agent. Otherwise, the project would not be installed at all.

(1985) "fundamental transformation"). For instance, hiring a new agent for the ongoing project might require specific training, which makes replacement unprofitable. See Spear and Wang (2005) and Mylovanov and Schmitz (2007) for dynamic contracting models in which replacement involves no costs.

[10] The assumption $I_2 \leq I_1$ is made only to simplify the presentation. If $I_2 > I_1$, then the project might be terminated even after a success in the first period, which would lead to more case distinctions without yielding any additional economic insights.

Contracts when effort is unobservable. In the remainder of the paper, we assume again that effort levels are unobservable. If the principal decides to install the project ($x_1 = 1$), a contract can specify the continuation decision $x_2(y_1)$ and transfer payments $t(y_1, y_2) \geq 0$ from the principal to the agent. Note that we do not impose any ad hoc restrictions on the class of feasible contracts; i.e., there is complete contracting in the sense of Tirole (1999).[11] It will be helpful to distinguish between (direct) rewards regarding the first and second period, so we use the following notational convention:

$$\begin{aligned} t(0,0) &= t_0 + t_{00}, \quad t(0,1) = t_0 + t_{01}, \\ t(1,0) &= t_1 + t_{10}, \quad t(1,1) = t_1 + t_{11}. \end{aligned}$$

It is easy to see that it is optimal for the principal to set $t_0 = t_{00} = t_{10} = 0$.[12] A first-period success will thus be directly rewarded with a bonus payment t_1, while a second-period success will be rewarded with a bonus t_{01} (following a first-period failure) or t_{11} (following a first-period success). As we will see, a first-period success will also be indirectly rewarded by the prospect of getting a larger bonus for a second-period success if it follows a first-period success, which will be a driving force behind our main results.

[11] As is well-known, in a pure moral hazard framework with unobservable effort, there is no need to consider messages sent by the agent (in contrast, such messages would have to be considered if effort were observable yet unverifiable). Moreover, note that there is no need to consider additional payments made at earlier dates. The principal is (weakly) better off when she makes all payments to the agent at date 4 (since the agent might "consume" earlier payments, so that he cannot be forced to pay them back).

[12] Obviously, the principal will set $t(0,0) = 0$. Hence, we can write $t_0 = t_{00} = 0$ and $t(0,1) = t_{01}$ without loss of generality. Moreover, note that we can restrict attention to $t_{11} \geq t_{10}$. If t_{10} were not zero, the payments could be replaced by $\tilde{t}_{10} = 0$, $\tilde{t}_1 = t_1 + t_{10}$, $\tilde{t}_{11} = t_{11} - t_{10}$. Notice that we could also assume that the direct rewards for the first period are paid at date 2 already.

3 Commitment: The "hot hand" effect

Let us first assume that the principal can commit not to renegotiate the contract that is written at date 0. If the principal installs the project ($x_1 = 1$), her expected profit is given by

$$\begin{aligned} & \Pi(e_1, e_2(0), e_2(1)) \\ = \; & e_1 \left(R - t_1 + x_2(1)\left[e_2(1)(R - t_{11}) - I_2\right]\right) \\ & +(1-e_1)x_2(0)\left[e_2(0)(R - t_{01}) - I_2\right] - I_1, \end{aligned}$$

otherwise it is given by zero. The agent's incentive compatibility constraints for the second period read

$$\begin{aligned} e_2(0) &= \arg\max_e et_{01} - c(e), \\ e_2(1) &= \arg\max_e et_{11} - c(e). \end{aligned}$$

The agent's first-period incentive compatibility constraint is given by

$$e_1 = \arg\max_e e\left[t_1 + e_2(1)t_{11} - c(e_2(1))\right] + (1-e)\left[e_2(0)t_{01} - c(e_2(0))\right] - c(e_1).$$

Moreover, the payments must be non-negative due to the agent's wealth constraint, which together with incentive compatibility also ensures participation.

The following proposition characterizes the second-best solution of the two-period model under full commitment.[13]

Proposition 1 *Assume that the principal can commit not to renegotiate. There exist a cut-off level I_2^* and continuous and decreasing threshold functions $I_1^C(I_2)$ and $I_1^T(I_2)$, where $0 < I_2^* < I_1^C(I_2^*) = I_1^T(I_2^*)$.*

a) If $I_2 \le I_2^*$ and $I_1 \le I_1^C(I_2)$, then the project is installed and it is always continued, $x_1 = x_2(0) = x_2(1) = 1$. The effort levels induced by the principal's optimal contract satisfy

$$e^{FB} \ge e_2^C(1) > e_1^C > e^{SB} > e_2^C(0) > 0.$$

[13]The superscript "C" denotes continuation, while "T" denotes termination.

b) If $I_2 > I_2^*$ and $I_1 \leq I_1^T(I_2)$, then the project is installed and it is terminated whenever the first period was a failure, $x_1 = x_2(1) = 1$, $x_2(0) = 0$. In this case,

$$e^{FB} \geq e_2^T(1) > e_1^T > e^{SB}.$$

c) Otherwise, the project is not installed at all, $x_1 = 0$.

Proof. See the appendix. ■

There are two ways in which a success in the first period is indirectly rewarded by the principal. First, consider the case in which the principal's investment costs I_1 and I_2 are sufficiently small, so that the project is installed and always continued (Proposition 1a). Even though a success in the first period has no technological effect whatsoever on the likelihood of a success in the second period, the principal implements $e_2^C(1) > e^{SB} > e_2^C(0)$. Giving the agent in the second period particularly high incentives following a first-period success (and particularly low incentives following a failure) has desirable spillover effects on the first-period incentives: The agent works hard in the first period not only in order to get the direct reward t_1, but also in order to enjoy a higher second-period rent ($A(e_2^C(1))$ instead of $A(e_2^C(0))$). Interestingly, from an outsider's perspective, this means that a success in the second period is indeed more likely to be observed after a first-period success; i.e., there is a "hot hand" effect endogenously generated purely due to incentive considerations.[14] In fact, the direct first-period reward t_1 will be positive only if the principal already induces $e_2^C(1) = e^{FB}$, so that implementing an even higher effort level following a first-period success would reduce the total surplus. Since giving the agent incentives in the first period is now cheaper than in the one-shot problem, the principal implements $e_1^C > e^{SB}$.

Second, the fact that the principal induces only low second-period effort after a first-period failure implies that continuing the project after a first-period failure

[14] See also McFall et al. (2006), who show a "hot hand" effect in a series of contests when there happens to be a huge reward for those who win the most contests. Players with initial good luck have more to win later on, therefore they exert relatively higher effort.

might not be in the principal's interest when her continuation costs I_2 are sufficiently large (Proposition 1b).[15] Clearly, if $e_2^C(0)$ is so small that $P(e_2^C(0)) < I_2$, the principal is better off if she terminates the project. But even if this inequality does not hold, it can still be optimal for the principal to commit to terminate the project, because doing so improves the agent's first-period incentives. Hence, the cut-off level I_2^* is smaller than $P(e_2^C(0))$.

The inefficiencies exhibited by the second-best solution are of a similar nature as the inefficiencies we encountered in the one-shot model. There are downward distortions of the effort levels compared with the first-best solution, and as a result there are projects that would be installed (and continued) in a first-best world, but that are not pursued (or at least not continued after a first-period failure) in the presence of moral hazard. However, it is still impossible for a project to be "too good" to be pursued, as is stated in the following corollary.

Corollary 1 *Assume that the principal can commit not to renegotiate. If at date 0 the principal can choose between two (otherwise identical) projects with possible returns $R = R_g$ and $R = R_b < R_g$, she will never prefer the bad project that can yield R_b only.*

Proof. See the appendix. ■

4 Renegotiation: The "overqualification" effect

After the first period is over, the principal might want to modify the contractual arrangements, because at that point in time she would be best off under the optimal one-period contract as characterized in Section 2.1. In the following we assume that the principal cannot ex ante commit not to renegotiate the contract. In our complete contracting framework, the principal can mimic the outcome

[15] Obviously, given that the principal installed the project at all, she will always continue if the first period was successful, because $I_2 \leq I_1$.

of renegotiations in her original contract; i.e., we can confine our attention to renegotiation-proof contracts.[16]

Proposition 2 *Assume that the principal cannot commit not to renegotiate. There exist continuous and decreasing threshold functions* $\bar{I}_1^C(I_2)$ *and* $I_1^T(I_2)$*, where* $P(e^{SB}) < \bar{I}_1^C(P(e^{SB})) < I_1^T(P(e^{SB}))$.

a) If $I_2 \leq P(e^{SB})$ *and* $I_1 \leq \bar{I}_1^C(I_2)$*, then the project is installed and it is always continued,* $x_1 = x_2(0) = x_2(1) = 1$. *The effort levels satisfy*

$$e^{FB} \geq \bar{e}_2^C(1) > \bar{e}_1^C > \bar{e}_2^C(0) = e^{SB}.$$

b) If $I_2 > P(e^{SB})$ *and* $I_1 \leq I_1^T(I_2)$*, then the project is installed and it is terminated whenever the first period was a failure,* $x_1 = x_2(1) = 1$, $x_2(0) = 0$, *and*

$$e^{FB} \geq e_2^T(1) > e_1^T > e^{SB}.$$

c) Otherwise, the project is not installed at all, $x_1 = 0$.

Proof. See the appendix. ■

As we have seen in the previous section, if the project was always continued under full commitment, the principal implemented a second-period effort level smaller than e^{SB} when the first period was a failure. The resulting smaller second-period rent acted as an indirect punishment of the wealth-constrained agent for the first-period failure. This is no longer possible if renegotiation cannot be ruled out, because at date 2 the principal would prefer to implement e^{SB} in order to maximize her second-period profit. While thus the "stick" is no longer available, the principal can still make use of the "carrot;" i.e., she can indirectly reward first-period effort by implementing an effort level larger than e^{SB} following

[16] Note that, in particular, this means that it is inconsequential how the renegotiation surplus would be split at date 2. The principal can achieve the same outcome that would be attained if she had all bargaining power in the renegotiation game by designing the appropriate renegotiation-proof contract at the outset.

a first-period success.[17] As a result, it is still cheaper for the principal to motivate the agent to exert first-period effort in the two-period model than in the one-shot benchmark model, so that $\bar{e}_1^C > e^{SB}$.

Just as in the full commitment regime, for sufficiently large investment costs I_2, the principal would be better off if she terminated the project whenever the first-period was a failure. This is clearly the case if continuation would cause more costs than gains, $I_2 > P(\bar{e}_2^C(0)) = P(e^{SB})$, but it is also the case for smaller investment costs I_2, because the threat to terminate improves the first-period incentives. Hence, there is a cut-off level $\bar{I}_2^* < I_2^* < P(e^{SB})$, in analogy to the commitment scenario. However, if renegotiation cannot be ruled out, at date 2 the principal prefers to continue the project as long as she can make a positive second-period profit by doing so. Her threat to terminate the project after a first-period failure is no longer credible, unless her expected second-period profit in case of continuation would actually be negative, $P(e^{SB}) - I_2 < 0$.

In other words, if $\bar{I}_2^* < I_2 < P(e^{SB})$, the principal would like to commit to termination following a first-period failure, but she cannot do so. This observation has peculiar implications with regard to the project that the principal will choose at the outset, as is highlighted in Corollary 3 below. A new kind of inefficiency occurs, which we saw neither in the well-known one-shot problem nor in the two-period model with full commitment.

Corollary 2 *Assume that the principal cannot commit not to renegotiate and $I_1 < \bar{I}_1^C(P(e^{SB}))$. If, ceteris paribus, I_2 is increased, then the principal's expected profit is reduced, except for the point $I_2 = P(e^{SB})$. At this point, there is an upward jump, which is bounded from below by $e^{SB}A(e^{SB})$.*

Proof. See the appendix. ■

[17] Note that the principal would like to reduce her promised payment t_{11} after a first-period success has occurred (in order to implement e^{SB} in the second period), but in this case there is no scope for mutually beneficial renegotiation. The agent would insist on the original contract, which gives him a larger rent.

Corollary 2 says that the principal can be better off if her continuation costs I_2 are increased, which may be surprising at first sight. Yet, this result follows immediately from the fact that the principal would like to commit to termination after a first-period failure for all $I_2 \geq \bar{I}_2^*$, but given that renegotiation cannot be ruled out, she can do so only when $I_2 \geq P(e^{SB})$. Hence, her expected profit makes an upward jump at $I_2 = P(e^{SB})$. This effect can be so strong that the principal would even prefer to have higher investment costs in both periods, or similarly, she would prefer to install a project that can only yield a smaller revenue R.

Corollary 3 *Assume that the principal cannot commit not to renegotiate. If at date 0 the principal can choose between two (otherwise identical) projects with possible returns $R = R_g$ and $R = R_b < R_g$, she may prefer the bad project that can yield R_b only.*

Proof. See the appendix. ■

For example, let $c(e) = \frac{1}{2}e^2$, $I_2 = 0.12$, $R_b = 0.68$, and $R_g = 0.7$. It is straightforward to show that the principal's expected profit is $\Pi \approx 0.147 - I_1$ if she installs the "good" project that can yield R_g, while it is $\Pi \approx 0.157 - I_1$ if she installs the "bad" project that can yield R_b only (and is otherwise identical). Note that if $I_1 = 0.15$, this even means that while the principal would be willing to install the "bad" project, the "good" project would never be funded.

Intuitively, pursuing a bad project that can yield a relatively small return (or, similarly, hiring a less qualified agent who can generate only a small return or who requires higher investments by the principal) acts as a commitment device. The principal knows that if she chooses the more attractive alternative, then at date 2 she cannot resist the temptation to continue after a first-period failure. For this reason, a project can be just "too good" to be funded or an "overqualified" agent may not be hired.[18]

[18] Lewis and Sappington (1993) have also pointed out that employers will sometimes not hire applicants who are "overqualified," even when their salary expectations are modest. However,

5 Concluding remarks

In this paper, we have extended the literature on repeated moral hazard problems to cover hidden action models in which the agent is risk-neutral but wealth-constrained. The present contribution seems to be sufficiently simple to be used as a building block in more applied work.

It is straightforward to relax several assumptions that were made to keep the exposition as clear as possible. For example, if it is required by an application, one might easily generalize the model by allowing different cost functions and different returns in the two periods. Moreover, one can dispense with the assumption that the principal has all bargaining power. Regardless of the bargaining protocol, the principal would only be willing to participate if her investment costs were covered. Hence, qualitatively our main findings would still be relevant. One could also consider the case in which the agent's wealth or his reservation utility may be positive. As long as his reservation utility is smaller than his rent, nothing changes. As long as the agent is not wealthy enough to "buy the firm," the effects highlighted in our model continue to be relevant.

their model is quite different from ours; they consider an adverse selection problem with countervailing incentives due to type-dependent reservation utilities. Note that in our model a more productive agent might not be hired even if his reservation utility is not higher than the one of a less qualified agent.

Appendix

Proof of Proposition 1.

The proof proceeds in several steps.

Step 1. Consider first the case $x_1 = x_2(0) = x_2(1) = 1$, so that the project is always continued. The incentive compatibility constraints can then be written as

$$\begin{aligned} c'(e_2(0)) &= t_{01}, \\ c'(e_2(1)) &= t_{11}, \\ c'(e_1) &= t_1 + A(e_2(1)) - A(e_2(0)). \end{aligned}$$

The principal chooses the payments or, equivalently, the effort levels to maximize her expected payoff

$$\Pi(e_1, e_2(0), e_2(1)) = e_1 \left[R - t_1 + P(e_2(1))\right] + (1 - e_1)P(e_2(0)) - I_2 - I_1,$$

where $t_1 = c'(e_1) - A(e_2(1)) + A(e_2(0))$ must be nonnegative. Thus, Π is a continuous function defined on the compact set $[0,1]^3 \cap \{t_1 \geq 0\}$ and as such it must have a maximum. The conditions on c ensure that the maximum is not on the boundary of $[0,1]^3$, but it could be on the boundary of $\{t_1 \geq 0\}$. The Kuhn-Tucker necessary condition for a solution of the maximization problem

$$\max_{e_1, e_2(0), e_2(1)} \Pi(e_1, e_2(0), e_2(1)) = P(e_1) + e_1 \left[S(e_2(1)) - S(e_2(0))\right] + P(e_2(0)) - I_2 - I_1$$

subject to $t_1 = c'(e_1) - A(e_2(1)) + A(e_2(0)) \geq 0$ is

$$\nabla \Pi + \lambda \nabla t_1 = 0 \quad \text{with } \lambda \geq 0 \text{ and } \lambda > 0 \Rightarrow t_1 = 0.$$

This gives us the three equations

$$\begin{aligned} P'(e_1^C) + S(e_2^C(1)) - S(e_2^C(0)) &= -\lambda c''(e_1^C), & (1) \\ e_1^C S'(e_2^C(1)) &= \lambda A'(e_2^C(1)), & (2) \\ P'(e_2^C(0)) - e_1^C S'(e_2^C(0)) &= -\lambda A'(e_2^C(0)). & (3) \end{aligned}$$

Furthermore, if $\lambda > 0$, then $t_1 = 0$ and using equation (1), we can calculate that

$$\lambda = e_1^C - \frac{R + P(e_2^C(1)) - P(e_2^C(0))}{c''(e_1^C)}.$$

Since P is bounded by $S(e^{FB}) < R$, this means that even if λ is positive, $\lambda < e_1^C$.

From equation (2), we see that

$$e_2^C(1) \le e^{FB} \text{ and } e_2^C(1) = e^{FB} \Leftrightarrow \lambda = 0.$$

This equation can also be written as

$$P'(e_2^C(1)) = \frac{\lambda - e_1^C}{e_1^C} A'(e_2^C(1)) < 0.$$

Thus, we can infer $e_2^C(1) > e^{SB}$. Likewise, equation (3) can be rearranged to

$$P'(e_2^C(0)) = \frac{e_1^C - \lambda}{1 - e_1^C} A'(e_2^C(0)) > 0.$$

Therefore, $e_2^C(0) < e^{SB}$. Moreover, equation (1) implies $e_1^C > e^{SB}$, because

$$P'(e_1^C) = S(e_2^C(0)) - S(e_2^C(1)) - \lambda c''(e_1^C) < 0.$$

Finally, $e_1^C < e_2^C(1)$ follows from

$$c'(e_1^C) = A(e_2^C(1)) - A(e_2^C(0)) < c'(e_2^C(1))$$

in the case $\lambda > 0$ and from

$$R - c'(e_1^C) = e_1^C c''(e_1^C) - S(e^{FB}) + S(e_2^C(0)) > t_1 + P(e_2^C(0)) > 0$$

(using the fact that $ec''(e) > c'(e)$ for $e > 0$) in the case $\lambda = 0$. Thus, we have proved that $e^{FB} \ge e_2^C(1) > e_1^C > e^{SB} > e_2^C(0) > 0$. The principal's expected profit in the case under consideration is

$$\Pi^C(I_1, I_2, R) = P(e_1^C) + e_1^C \left[S(e_2^C(1)) - S(e_2^C(0))\right] + P(e_2^C(0)) - I_2 - I_1.$$

Step 2. Consider next the case in which the principal chooses $x_1 = x_2(1) = 1$, $x_2(0) = 0$, so that the project is terminated whenever the first period was a failure. The optimal contract can then be characterized in analogy to Step 1. In particular, the first-order conditions now read

$$\begin{aligned} P'(e_1^T) + S(e_2^T(1)) - I_2 &= -\lambda c''(e_1^T), \\ e_1^T S'(e_2^T(1)) &= \lambda A'(e_2^T(1)), \end{aligned}$$

and $t_1 = c'(e_1^T) - A(e_2^T(1)) \geq 0$ is binding if $\lambda > 0$. In analogy to Step 1, it follows that $e^{FB} \geq e_2^T(1) > e_1^T > e^{SB}$.[19] The expected profit in this case is

$$\Pi^T(I_1, I_2, R) = P(e_1^T) + e_1^T \left[S(e_2^T(1)) - I_2\right] - I_1.$$

Step 3. It is straightforward to check that $x_1(1) = 0$, $x_2(0) = 1$ can never be optimal. Among the remaining alternatives (the cases discussed in Steps 1 and 2 and $x_1 = 0$), the principal will always commit to the one leading to the largest expected profits. For a given R, the expected profits $\Pi^C(I_1, I_2, R)$ and $\Pi^T(I_1, I_2, R)$ are decreasing in I_1 and I_2. In particular, using the envelope theorem it follows that

$$\frac{d\Pi^T(I_1, I_2, R)}{dI_2} = -e_1^T > -1 = \frac{d\Pi^C(I_1, I_2, R)}{dI_2}.$$

Note that $\Pi^C(I_1, 0, R) - \Pi^T(I_1, 0, R)$ does not depend on I_1 and is strictly positive, because in Step 1 we showed that it is not optimal to set $e_2^C(0) = 0$. Moreover, if I_2 were prohibitively large, continuation would never be optimal. Hence, there exists a unique cut-off level $I_2^* > 0$, such that $\Pi^C(I_1, I_2^*, R) = \Pi^T(I_1, I_2^*, R)$. If $I_2 \leq I_2^*$, continuation is better than termination, and if in addition $I_1 \leq I_1^C(I_2) := \Pi^C(0, I_2, R)$, then $x_1 = 1$ is optimal. If $I_2 > I_2^*$ and $I_1 \leq I_1^T(I_2) := \Pi^T(0, I_2, R)$, then the project is started and it is terminated following a first-period failure. Otherwise, the project is not installed. It remains to be shown that the region in which the termination contract is optimal is nonempty, given that $I_2 \leq I_1$. As has already been pointed out in the discussion

[19] Note that to show $e_1^T > e^{SB}$, we now use $S(e_2^T(1)) > I_2$, which due to $I_2 \leq I_1$ must hold if the case under consideration is more profitable for the principal than $x_1 = 0$.

following Proposition 1, $I_2^* < P(e_2^C(0))$.[20] Hence,

$$I_1^C(I_2^*) = \Pi^C(0,0,R) - I_2^* > 2P(e^{SB}) - P(e^{SB}) > I_2^*,$$

and the proposition has been proved. ■

Proof of Corollary 1.

Consider the optimal contract in the case of the bad project ($R = R_b$). In the case of the good project ($R = R_g$), the principal could simply offer the same contract. Then the agent's behavior would be the same, but the principal's expected profit would be larger. By optimally adjusting the contract in the case of the good project, the principal can do even better. ■

Proof of Proposition 2.

The proof proceeds again in several steps.

Step 1. The principal now has to take into consideration additional renegotiation-proofness constraints. First, if the project is continued, then $e_2(y_1) \in [e^{SB}, e^{FB}]$ must be satisfied. If the original contract induced $e_2(y_1) < e^{SB}$, then at date 2 both parties could be made better off if the transfer payments were renegotiated such that the agent chooses e^{SB}. Similarly, implementing $e_2(y_1) = e^{FB}$ would make both parties better off if the original contract induced $e_2(y_1) > e^{FB}$. If $e_2(y_1) \in [e^{SB}, e^{FB}]$, no renegotiation occurs, because effort levels closer to e^{FB} are more efficient, but given that transfer payments to the agent must be non-negative, the principal's second-period profit is larger the closer the induced effort level is to e^{SB} (see Section 2.1). Second, if the project is installed and $I_2 \leq P(e^{SB})$, then $x_2(y_1) = 1$ must be satisfied (if $x_2(y_1) = 0$, renegotiation would occur, because at date 2 both parties could be made better off by continuing the project). If $I_2 > P(e^{SB})$, then $x_2(y_1) = 0$ is renegotiation-proof, whereas $x_2(y_1) = 1$ is renegotiation-proof whenever $I_2 \leq S(e_2(y_1))$.

[20] To see this formally, note that $\Pi^C(I_1, P(e_2^C(0)), R) < P(e_1^C) + e_1^C \left[S(e_2^C(1)) - P(e_2^C(0))\right] - I_1 < \Pi^T(I_1, P(e_2^C(0)), R)$.

Step 2. Consider the case $I_2 \leq P(e^{SB})$. Given that the project is installed, the analysis is analogous to Step 1 of the proof of Proposition 1, where $x_1 = x_2(0) = x_2(1) = 1$. In particular, equation (3) now has to be replaced with $\bar{e}_2^C(0) = e^{SB}$, since the renegotiation-proofness constraint $\bar{e}_2^C(0) \geq e^{SB}$ is always binding (among all renegotiation-proof effort levels, e^{SB} not only maximizes the principal's second-period expected profit, but it also provides highest incentives for the agent in the first period). The other two effort levels, $\bar{e}_1^C$ and $\bar{e}_2^C(1)$, have to be adjusted accordingly. They can be derived from the first order conditions (1) and (2), and the implications for the ordering of the effort levels remain true. The principal's expected profit in the case of unconditional continuation is

$$\bar{\Pi}^C(I_1, I_2, R) = P(\bar{e}_1^C) + \bar{e}_1^C \left[S(\bar{e}_2^C(1)) - S(e^{SB})\right] + P(e^{SB}) - I_2 - I_1.$$

The project will not be installed ($x_1 = 0$) if $I_1 > \bar{I}_1^C(I_2) := \bar{\Pi}^C(0, I_2, R)$.

Step 3. Consider next the case $I_2 > P(e^{SB})$. Since then $I_2 > I_2^*$, the termination contract ($x_1 = x_2(1) = 1$, $x_2(0) = 0$) characterized in Step 2 of Proposition 1 solves the principal's maximization problem, given that $x_1 = 1$. This contract is renegotiation-proof, because $e_2^T(1) > e^{SB}$ and $I_2 \leq I_1 \leq P(e_1^T) + e_1^T \left[S(e_2^T(1)) - I_2\right] \leq S(e_2^T(1))$, given that the principal decides to install the project. The project will not be installed if $I_1 > I_1^T(I_2)$. Finally, note that

$$I_1^T(P(e^{SB})) > \bar{I}_1^C(P(e^{SB})) = \bar{\Pi}^C(0,0,R) - P(e^{SB}) > 2P(e^{SB}) - P(e^{SB}) = P(e^{SB}),$$

so that the proposition follows immediately. ■

Proof of Corollary 2.

$\bar{\Pi}^C(I_1, I_2, R)$ and $\Pi^T(I_1, I_2, R)$ are continuous and decreasing in I_2. At $I_2 = P(e^{SB}) > I_2^*$, we know that $\Pi^T(I_1, I_2, R) > \Pi^C(I_1, I_2, R) > \bar{\Pi}^C(I_1, I_2, R)$. Hence, given that the project is installed, at $I_2 = P(e^{SB})$ the principal's expected profit as characterized in Proposition 2 is discontinuous, and the size of the jump is given by

$$
\begin{aligned}
& \Pi^T(I_1, P(e^{SB}), R) - \bar{\Pi}^C(I_1, P(e^{SB}), R) \\
= \; & P(e_1^T) + e_1^T \left[S(e_2^T(1)) - P(e^{SB})\right] - \left(P(\bar{e}_1^C) + \bar{e}_1^C \left[S(\bar{e}_2^C(1)) - S(e^{SB})\right]\right) \\
> \; & P(\bar{e}_1^C) + \bar{e}_1^C \left[S(\bar{e}_2^C(1)) - P(e^{SB})\right] - \left(P(\bar{e}_1^C) + \bar{e}_1^C \left[S(\bar{e}_2^C(1)) - S(e^{SB})\right]\right) \\
> \; & e^{SB} A(e^{SB}).
\end{aligned}
$$

■

Proof of Corollary 3.

Take any $R = \tilde{R}$ and I_1, I_2 such that $I_2 = P(e^{SB})$ and $I_1 \leq I_1^T(I_2)$. Corollary 2 shows that $\Pi^T(I_1, I_2, \tilde{R}) > \bar{\Pi}^C(I_1, I_2, \tilde{R}) + e^{SB} A(e^{SB})$. Note that, using the envelope theorem, $dP(e^{SB})/dR = e^{SB} > 0$. Proposition 2 thus implies that if R_g is slightly larger than $\tilde{R}$ and R_b is slightly smaller than $\tilde{R}$, then the principal prefers R_b (leading to the expected profit $\Pi^T(I_1, I_2, R_b)$) to R_g (leading to the expected profit $\bar{\Pi}^C(I_1, I_2, R_g)$). ■

References

Baliga, S. and Sjöström, T. "Decentralization and Collusion." *Journal of Economic Theory,* Vol. 83 (1998), pp. 196–232.

Bolton, P. and Dewatripont, M. *Contract Theory.* Cambridge, MA: MIT Press, 2005.

Che, Y.-K. and Yoo, S.-W. "Optimal Incentives for Teams." *American Economic Review,* Vol. 91 (2001), pp. 525–541.

Chiappori, P., Macho, I., Rey, P., and Salanié, B. "Repeated Moral Hazard: The Role of Memory Commitment and the Access to Credit Markets." *European Economic Review,* Vol. 38 (1994), pp. 1527–1553.

Crémer, J. "Arm's Length Relationships." *Quarterly Journal of Economics,* Vol. 110 (1995), pp. 275–295.

Dewatripont, M., Legros, P., and Matthews, S. "Moral Hazard and Capital Structure Dynamics." *Journal of the European Economic Association,* Vol. 1 (2003), pp. 890–930.

Dewatripont, M. and Maskin, E. "Credit and Efficiency in Centralized and Decentralized Economies." *Review of Economic Studies,* Vol. 62 (1995), pp. 541–555.

Fudenberg, D., Holmström, B., and Milgrom, P. "Short-Term Contracts and Long-Term Agency Relationships." *Journal of Economic Theory,* Vol. 51 (1990), pp. 1–31.

Fudenberg, D. and Tirole, J. "Moral Hazard and Renegotiation in Agency Contracts." *Econometrica*, Vol. 58 (1990), pp. 1279–1319.

Gigler, F. and Hemmer, T. "On the Welfare Effects of Allowing Unlimited Renegotiation in Agency Relationships." *Economic Theory,* forthcoming, 2007.

Gilovich, T., Vallone, R., and Tversky, A. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology*, Vol. 17 (1985), pp. 295–314.

Hermalin, B.E. and Katz, M.L. "Moral Hazard and Verifiability: The Effects of Renegotiation in Agency." *Econometrica*, Vol. 59 (1991), pp. 1735–1753.

Hirao, Y. "Learning and Incentive Problems in Repeated Partnerships." *International Economic Review*, Vol. 34 (1993), pp. 101–119.

Innes, R.D. "Limited Liability and Incentive Contracting with Ex-ante Action Choices." *Journal of Economic Theory,* Vol. 52 (1990), pp. 45–67.

Kornai, J., Maskin, E., and Roland, G. "Understanding the Soft Budget Constraint." *Journal of Economic Literature,* Vol. 41 (2003), pp. 1095–1136.

Laffont, J.-J., and Martimort, D. *The Theory of Incentives: The Principal-Agent Model.* Princeton, N.J.: Princeton University Press, 2002.

Lazear, E.P. "Agency, Earnings Profiles, Productivity, and Hours Restrictions." *American Economic Review,* Vol. 71 (1981), pp. 606–620.

Lewis, T.R. and Sappington, D.E.M. "Choosing Workers' Qualifications: No Experience Necessary?" *International Economic Review,* Vol. 34 (1993), pp. 479–502.

Ma, C.-T. A. "Adverse Selection in Dynamic Moral Hazard." *Quarterly Journal of Economics,* Vol. 106 (1991), pp. 255–275.

Ma, C.-T. A. "Renegotiation and Optimality in Agency Contracts." *Review of Economic Studies,* Vol. 61 (1994), pp. 109–129.

Malcomson, J.M. and Spinnewyn, F. "The Multiperiod Principal-Agent Problem." *Review of Economic Studies,* Vol. 55 (1988), pp. 391–407.

Matthews, S. "Renegotiation of Sales Contracts." *Econometrica,* Vol. 63 (1995), pp. 567–591.

Matthews, S. "Renegotiation of Moral Hazard Contracts under Limited Liability and Monotonicity." *Journal of Economic Theory,* Vol. 97 (2001), pp. 1–29.

McFall, T.A., Knoeber, C.R., and Thurman, W.N. "Contest, Grand Prizes, and the Hot Hand." Discussion Paper, 2006.

Mukoyama, T. and Şahin, A. "Repeated Moral Hazard with Persistence." *Economic Theory,* Vol. 25 (2005), pp. 831–854.

Mylovanov, T. and Schmitz, P.W. "Task Scheduling and Moral Hazard." *Economic Theory,* forthcoming, 2007.

Pitchford, R. "Moral Hazard and Limited Liability: The Real Effects of Contract Bargaining." *Economics Letters,* Vol. 61 (1998), pp. 251–259.

Rey, P. and Salanié, B. "Long-Term, Short-Term and Renegotiation: On the Value of Commitment in Contracting." *Econometrica,* Vol. 58 (1990), pp. 597–619.

Rogerson, W.P. "Repeated Moral Hazard." *Econometrica,* Vol. 53 (1985), pp. 69–76.

Schmitz, P.W. "Allocating Control in Agency Problems with Limited Liability and Sequential Hidden Actions." *RAND Journal of Economics,* Vol. 36 (2005), pp. 318–336.

Shapiro, C. and Stiglitz, J. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review,* Vol. 74 (1984), pp. 433–444.

Spear, S.E. and C. Wang (2005). "When to Fire a CEO: Optimal Termination in Dynamic Contracts." *Journal of Economic Theory,* Vol. 120 (2005), pp. 239–256.

Tamada, Y. and Tsai, T.-S. "Optimal Organization in a Sequential Investment Problem with the Principal's Cancellation Option." *International Journal of Industrial Organization,* Vol. 25 (2007), pp. 631–641.

Tirole, J. "Incomplete Contracts: Where Do We Stand?" *Econometrica,* Vol. 67 (1999), pp. 741–781.

Tirole, J. "Corporate Governance." *Econometrica,* Vol. 69 (2001), pp. 1–35.

Wang, C. "Renegotiation-Proof Dynamic Contracts with Private Information." *Review of Economic Dynamics,* Vol. 3 (2000), pp. 396–422.

Williamson, O.E. *The Economic Institutions of Capitalism.* New York: The Free Press, 1985.