Contractual contingencies and renegotiation

M. Dewatripont*

and

E. Maskin**

In a dynamic model of asymmetric information between the owner of a firm and a manager, we investigate the optimal set of contingencies on which an incentive contract should depend when renegotiation is possible. In particular, we characterize the circumstances in which the contracting parties find it desirable to deliberately restrict what the owner can monitor, thereby limiting the contractible contingencies. Our findings thus provide an endogenous explanation for contract simplicity, in contrast to those based on transactions costs or bounded rationality.

1. Introduction

The issue of contract renegotiation has recently attracted some attention in the incentives literature. See, for example, Dewatripont (1988, 1989), Dewatripont and Maskin (1990), Hart and Tirole (1988), Laffont and Tirole (1990), Fudenberg and Tirole (1990), Hermelin and Katz (1991), and Segal and Tadelis (1994). All these articles, as well as our own, are concerned with renegotiation that arises in models of asymmetric information; see Dewatripont and Maskin (1990) for a survey. The literature to date, however, has concentrated on models where there is only a single screening variable (by a “screening variable” we mean a publicly observable decision variable under the control of a party with private information). In this article, in contrast, we allow for multiple variables. This enables us to investigate the optimal set of screening variables. Specifically, we ask when increasing the number of publicly observable variables is a good thing or when restricting observability is advantageous.

---

* DUBLEA and ECARE, Université Libre de Bruxelles, CEPR, and CORE.
** Harvard University.

We thank Louis Makowski, Margaret Meyer, Patrick Rey, Michael Riordan, Michael Whinston, and two referees for helpful comments, Steven Tadelis for research assistance, and the National Science Foundation and the ARC-EST research program for research funding.

1 Another body of work (e.g., Hart and Moore (1988), Green and Laffont (1992), Aghion, Dewatripont, and Rey (1994), and Maskin and Moore (1987) examines renegotiation when parties have the same information. See also the literature on renegotiation in repeated games, e.g., Bernheim and Ray (1989), Farrell and Maskin (1989), and Pearce (1987).

2 There is a small literature on multiple screening variables, including Maskin and Riley (1985) and Matthews and Moore (1987).
Were the parties able to commit themselves not to renegotiate, increasing the number of observable variables would necessarily be desirable (at least weakly); clearly, a contract can only be improved by expanding the set of observables on which it can be made contingent. This implication does not follow, however, if renegotiation is possible. Obviously, one observable variable is better than none, but two variables set in sequence may result in information being revealed too quickly: the level at which the first variable is set may convey information about the informed party’s type, and because of renegotiation, this can interfere with the second variable’s risk-sharing role. If interference is severe enough, it may be optimal for the first variable to be left completely unobservable by the uninformed party. Intuitively, this is likely when the second variable is a “better” screening device than the first (so that the loss in screening power by making the first unobservable is more than compensated for by enhancing the value of the second variable) and when it is “costly” to avoid having the first variable reveal information, assuming that it is observable. One of the main tasks of this article is to make precise these intuitive but vague ideas.

Our results explain how contracts with fewer contingencies can dominate more complex contracts when the various screening variables end up “conflicting” with one another. Hence, one can interpret the article as motivating the use of simple, relatively uncontingent contracts in adverse-selection problems. A question arises, however: Can parties credibly restrict observability? The answer could well be “yes” if observability requires setting up some monitoring procedure in advance, say, at the time the contract is signed. Suppose, for example, that the variable in question is intangible capital. If the monitoring procedure consists of sending somebody to check how much is being invested, failing to send this monitor at the outset may make acquiring meaningful information impossible afterward. In such a case, our results suggest that even if it is costless to send the monitor in advance, it might be rational not to do so, so as to restrict observability ex post. Thus, our explanation of the superiority of simple contracts, unlike an informal story that is often told, does not rely on the cost of enforcing or enumerating contingencies.

We investigate a model in which a risk-averse manager contracts with the owner of a firm to perform a specific task. There is uncertainty ex ante about the extent to which the manager can reduce his own (unobservable) effort into this task by investing in capital and hiring subordinates (which are, in principle, observable to the owner). After signing the contract the manager—but not the owner—learns the resolution of the uncertainty. He then invests in capital and finally hires subordinates. Renegotiation can occur between the capital- and labor-setting stages. The contract provides the manager with a budget out of which he can hire capital and labor (he retains the residual for himself). Since the manager is risk averse, this budget should ideally depend on the realization of the uncertainty, but because of the asymmetry of information, it can do so only indirectly through the potentially observable variables, capital and labor. A two-variable contract specifies the manager’s budget as a function of both the capital and employment levels. We compare such a contract with the scheme that results when the owner observes only one factor: a one-variable contract. The comparison is made in two steps.

In Section 2 we restrict attention to one-variable contracts and derive the “cheap factor principle” (Proposition 1): If the owner can monitor only one input and there is sufficient uncertainty, the contracting parties are better off if the monitored input is the cheaper of the two. We then apply this principle in Section 3 to show that when labor is sufficiently cheap relative to capital (making labor the better screening device), one-variable contracts may dominate those with two variables (Proposition 3).
In Section 4 we alter the informational structure of our model and assume that the owner can, in principle, observe the cost saving created by capital and labor but not the labor input level itself. We then compare contracts with and without capital observability. In this framework, the counterpart of the question considered in Proposition 1 is whether it is better to make contracts contingent on capital or on cost saving. Maskin and Riley (1985) gave a general answer to this question. Translated into our framework, they showed that under plausible hypotheses, the cost-saving contract dominates. This result suggests that a one-variable (cost-saving) contract may be preferable to two variables if renegotiation is possible, since that way full scope is given to cost-saving as a screening device. Indeed, as we confirm in Proposition 4, the one-variable contract dominates under circumstances quite analogous to those of Proposition 3. Section 5 offers some concluding remarks.

There are several other articles that address related issues. Riordan (1987) examines a model of procurement in which the government must decide whether or not to observe a signal about production costs before offering a defense firm a contract. The government may refrain from doing so because the signal will affect the contract it chooses to offer the firm and therefore the firm’s incentive to undertake investment beforehand (see also Riordan (1990)). Riordan, unlike us, rules out the possibility of an ex ante contract. Crémer (1995) considers a repeated principal-agent model in which the agent has to choose an effort level twice and where the principal has the possibility of learning information about the agent’s intrinsic ability after her first effort choice. Crémer shows that not observing this information may make the principal more credible in punishing the agent after poor first-period performance, thereby inducing higher first-period effort. Finally, Dewatripont and Maskin (forthcoming) show that it may be beneficial for a creditor to restrict his observability of an entrepreneur’s activity because this may make the threat of not refinancing bad projects credible and thereby deter poor entrepreneurs from asking for funds in the first place.

Whereas all the above articles explain how lack of observability may improve commitment by a single party, the current article focuses on the interaction between observability and Pareto-improving renegotiation.

2. Single input monitoring: the “cheap factor principle”

Before looking at sequential screening and renegotiation, we first consider the determinants of efficient screening when only one variable is available. If the manager employed neither capital nor labor, the cost he would incur (measured in money) from performing the task is assumed to be $E$, which can be thought of, for example, as his disutility of effort or as the imputed cost of his time. By setting capital and labor equal to $k$ and $l$, respectively, he can achieve a cost saving of $\theta f(k, l)$ (which can be interpreted as the time he saves by delegating), where $\theta$ is the realization of a shock. This shock, which is observable only to the manager, may be viewed as a measure of how readily he can delegate part of the task to others. We assume that $\theta$ takes on only two values, $\theta_1$ and $\theta_2$, $0 < \theta_1 < \theta_2$, with probabilities $p_1$ and $p_2 = 1 - p_1$. Hence, $\theta_2$ corresponds to the case where delegation is easy, and $\theta_1$ to where it is hard. We shall assume throughout that $f$ is twice differentiable and concave on the nonnegative orthant, that $f_j > 0$ for $j = 1, 2$ ($f_j$ is the first derivative of $f$ with respect to its $j$th argument), and that $f_{12} \geq 0$ (where $f_{12}$ is the cross-partial derivative; a stronger hypothesis is needed for Propositions 3 and 4). For the time being we shall suppose that $f$ is symmetric, i.e., that $f(k, l) = f(l, k)$. Assume the unit costs of labor and capital to be $w$ and $r$ respectively.

---

3 We shall drop this assumption of symmetry in subsequent sections.
Let $V$ be the manager’s von Neumann-Morgenstern utility function, where $V'' < 0$ (primes denote derivatives). We shall assume that $V$ satisfies the following marginal utility ratio condition (MURC):

$$\lim_{x \to \infty, y \to \infty} V'(x + y)/V'(x) = 0.$$ 

(Note that the limit on the left-hand side is required to exist—and equal zero—for any sequences of $x$'s and $y$'s tending to infinity.) MURC is satisfied by utility functions that, for example, exhibit constant absolute risk aversion⁴ or infinite risk aversion, and it holds approximately for extreme risk aversion.

The manager, being risk averse, wishes to be insured against the variability of $\theta$ by the owner, who we suppose is risk neutral. The manager learns the value of $\theta$ before setting $k$ and $l$, but the owner never observes $\theta$ directly.

Let us assume, for convenience, that the manager has all the bargaining power in the design of the incentive constraint. Then if $I$ is the maximum expected amount that the owner will pay to have the task performed, and if labor is the only variable that can be monitored, the optimal incentive contract signed before the manager learns the value of $\theta$ solves

$$\max_{k_i, l_i, I_i} \sum_{i=1}^{2} p_i V\left(\theta f(k_i, l_i) - E - r k_i - w l_i + I_i\right)$$

subject to

$$\sum_{i=1}^{2} p_i I_i \leq I,$$

for $i, j = 1, 2$ and $i \neq j,$

and

$$\begin{cases} 
\{\theta f_i(k_i, l_i) \leq r \} 
\cap 
\{k_i \geq 0 \} 
\end{cases}$$

for $i, j = 1, 2$ and $i \neq j,$

where $l_i, k_i,$ and $I_i$ are respectively the employment level, capital investment, and budget in state $\theta_i$, $k_{ij}$ is the capital level a type-$i$ firm will set when "pretending" to be type $j$, and each embraced pair in (4) is complementarily slack. Let $(l^0, k^0, k^0_{ij})$ be the solution to this program.⁵

⁴ In that case, $V'(x + y)/V'(x) = e^{-r(x+y)}/e^{-r} = e^{-r}$, which goes to 0 as $y \to \infty$.

⁵ Note that in the program (1)–(4) we are ignoring the possibility that $l_i$ and $k_i$ might be random variables. In fact, it can be shown (see Maskin (1981)) that random contracts are in general desirable. Because considering such contracts would not qualitatively affect our results, however, we rule them out to keep the analysis as simple as possible.
In the above formulation, \( \theta f(k_i, l_i) - E - rk_i - wl_i + I_i \) represents the manager’s net benefit in state \( \theta_i \), and formula (1) is his expected utility. Inequality (2) is the owner’s individual rationality constraint. Inequalities (3) are the manager’s incentive (self-selection) constraints. In fact, at the optimum, only the incentive constraint for \( i = 2 \) will be binding, i.e., the one that constrains the manager from falsely claiming \( \theta = \theta_1 \). Since only labor is monitored, the manager has to set \( i = l_i \) in order to receive \( I_i \), but he can set capital at the efficient level given \( l_i \), which implies the inequalities (4). The capital level \( k_{ij} \) thus maximizes net benefit given that \( \theta = \theta_i \) and \( l = l_i \). Note that the maximization of (1) subject to (2) and (3) implies that the second embraced pair in (4) holds automatically. Moreover, the variable \( k_{12} \) appears only in the irrelevant incentive constraint.

Henceforth, let us normalize payoffs so that \( E = I = 0 \). Then the optimal contract \( (l_i^{00}, k_i^{00}, 0_2^{00}) \) when only capital can be monitored solves

\[
\max_{k_i, l_i, l_i} \sum_{i=1}^{2} p_i V(\theta f(k_i, l_i) - rk_i - wl_i + I_i)
\]

subject to

\[
\sum_{i=1}^{2} p_i l_i = 0, \quad \theta_2 f(k_2, l_2) - rk_2 - wl_2 + I_2 \geq \theta_2 f(k_2, l_2) - rk_2 - wl_2 + I_2, \quad \text{capital monitored}
\]

and

\[
\begin{cases}
\theta_2 f(k_1, l_2) \leq w \\
l_2 \geq 0
\end{cases}
\quad \text{and} \quad \begin{cases}
\theta_2 f(k_i, l_i) \leq w \\
l_i \geq 0
\end{cases}
\quad \text{for } i = 1, 2,
\]

where \( l_2 \) is the profit-maximizing employment level given \( \theta = \theta_2 \) and \( k = k_1 \), and each embraced pair in (4a) is complementarily slack. We are now ready to state the “cheap factor principle”.

**Proposition 1.** If \( r > w \), then for \( \theta_2 \) big enough relative to \( \theta_1 \), the manager is better off when labor rather than capital is monitored. The opposite conclusion holds if \( r < w \).

**Proof.** See the Appendix.

The cheap factor principle asserts that when there is sufficient uncertainty, it is better to monitor the cheaper rather than the more expensive of two factors. The economic reasoning behind this result is fairly straightforward. On the one hand, the standard underemployment feature of optimal allocations under adverse selection implies that a manager in state \( \theta_i \) uses the monitored input less than the unmonitored one (relative to input costs). This consideration favors monitoring the expensive factor, since it implies that, in that case, the manager’s payoff in state \( \theta_i \) will be higher (exclusive of transfers from the owner) than it would be were the cheap input monitored. (If \( \theta_2 \)

---

6 The other constraint, preventing a false claim that \( \theta = \theta_2 \), will not be binding. This is so because if the manager did falsely claim that \( \theta = \theta_2 \), the owner would actually be better off (since it is easy to show that \( l_1 > l_2 \)). Therefore, there is no need to constrain the manager when \( \theta = \theta_1 \) (see Hart (1983) for details).
is big enough relative to \( \theta_i \), then the manager’s payoff in state \( \theta_i \)—excluding transfers—is approximately \( \theta_i f(k_0^i, 0) - rk_1^0 \) when labor is monitored, whereas it is nearly

\[
\theta_i f(0, l_1^{00}) - wl_1^{00}
\]

when capital is monitored. Clearly, when \( r > w \), the latter is higher.) On the other hand, this difference in input use is also true of a \( \theta_2 \)-manager who falsely claims to be in state \( \theta_i \). That is, the cost of “cheating” by such a manager is higher if it is the expensive factor that is left unmonitored. Hence, when \( r > w \), monitoring labor rather than capital has the stronger influence on relaxing the incentive constraint. Moreover, for \( \theta_2 \) big enough, this relaxation effect dominates the above \( \theta_1 \)-profit effect, since for fixed levels of capital and labor, the difference between the right-hand and left-hand sides of the incentive constraints (3) and (3a) grows in \( \theta_2 \). Hence, with sufficient uncertainty, it is better to monitor labor. Put another way, an expensive factor “takes care of itself,” since it will be used in small quantities whether or not it is monitored. Monitoring the cheap factor thus has a greater influence on managerial behavior.

To see that the cheap factor principle applies only when there is sufficient uncertainty (i.e., \( \theta_2 \) big enough relative to \( \theta_i \)), it is instructive to examine the case in which \( \theta_2 \) is very close to \( \theta_1 \). To simplify computations, let us suppose that the manager has infinite risk aversion. This means that he cares only about his payoff in state \( \theta_i \). Hence, under labor monitoring his payoff is

\[
\theta_i f(k_1^0, l_1^{00}) - rk_1^0 - w l_1^{00} + p_2 \left( \theta_2 f(k_2^0, l_2^{00}) - rk_2^0 - w l_2^{00} - \theta_2 f(k_2^0, l_2^{00}) + rl_2^0 + wl_1^{00} \right),
\]

(5a)

whereas under capital monitoring it is

\[
\theta_i f(k_1^{00}, l_1^{00}) - rk_1^{00} - w l_1^{00}
\]

\[+ p_2 \left( \theta_2 f(k_2^{00}, l_2^{00}) - rk_2^{00} - w l_2^{00} - \theta_2 f(k_2^{00}, l_2^{00}) + rk_2^{00} + w l_1^{00} \right).
\]

(5b)

Treating all variables as functions of \( \theta_i \) and differentiating (5a) and (5b) with respect to \( \theta_i \) at the point \( \theta_i = \theta_2 \), we obtain

\[
f(k_1^0, l_1^0)
\]

and

\[
f(k_1^{00}, l_1^{00})
\]

respectively. (All other terms vanish thanks to the envelope theorem.) Differentiating again with respect to \( \theta_i \) at \( \theta_i = \theta_2 \), we obtain

\[
f_1(k_1^0, l_1^0)k_1^{00} + f_2(k_1^0, l_1^{00})l_1^{00}
\]

(7a)

and

\[
f_1(k_1^{00}, l_1^{00})k_1^{00} + f_2(k_1^{00}, l_1^{00})l_1^{00}
\]

(7b)
where primes denote derivatives with respect to $\theta_i$. The derivatives $k_i'$ and $l_i'$ can be computed by implicitly differentiating the first-order conditions $\theta_i f_1(k_i', l_i') = r$, $\theta_2 f_2(k_2', l_2') = r$, and $\theta_i f_2(k_i', l_i') - w - p_2(\theta_2 f_2(k_2', l_2') - w) = 0$. Hence we can rewrite (7a) as

$$2f_{12} f_2 - f_2^2 f_{11} - p_2 \frac{f_1^2 f_{12}}{f_{11}} - p_1 f_{22} f_1^{1/2}$$

$$\theta_2 p_1(f_{11} f_{22} - f_{12}^{1/2}).$$

Similarly, (7b) can be rewritten as

$$2f_{12} f_1 f_2 - f_2^2 f_{22} - p_2 \frac{f_2^2 f_{12}}{f_{22}} - p_1 f_1 f_2^{1/2}$$

$$\theta_1 p_1(f_{11} f_{22} - f_{12}^{1/2}).$$

Subtracting (8b) from (8a), we obtain

$$\frac{p_2}{\theta_2 p_1} \left( \frac{f_2^{1/2}}{f_{22}^{1/2}} - \frac{f_1^{1/2}}{f_{11}^{1/2}} \right).$$

Now if (9) is negative, then (5a) is greater than (5b) for $\theta_2$ near $\theta_1$, so the cheap factor principle applies when $r > w$. But if (9) is negative—which is entirely consistent with our assumptions about $f$, $r$, and $w$—we find that (5b) exceeds (5a) for $\theta_2$ near $\theta_1$, i.e., an “expensive factor principle” holds.

3. Single- versus multiple-input monitoring schemes

We now turn to contracts in which both inputs are monitored and the manager sets capital before labor. We suppose that after capital is in place the owner can propose a new contract, which takes effect if the manager agrees. If the manager rejects the proposal, the original contract remains in force. If there exists a (perfect Bayesian) equilibrium of this renegotiation game in which the original contract is not renegotiated, we say that it is renegotiation-proof.7

For convenience, we rule out renegotiation before capital is set. We can do so without loss of generality, since parties always choose Pareto-optimal contracts, and a contract that is Pareto optimal in the class of contracts that are renegotiation-proof after capital is set is also renegotiation-proof beforehand. (If $\langle l_i, k_i, l_i' \rangle$ is Pareto optimal in the class of ex post renegotiation-proof contracts but is not ex ante renegotiation-proof, then the owner gains from proposing some alternative contract $\langle l_i', k_i, l_i' \rangle$. But, without loss of generality, we can assume that the latter contract is ex post renegotiation-proof, since otherwise we can replace it with the renegotiated contract. Hence $\langle l_i', k_i, l_i' \rangle$ contradicts $\langle l_i, k_i, l_i' \rangle$’s Pareto optimality.)

The previous section considered only one-variable schemes, so that renegotiation could not arise. With sequential monitoring of both inputs, renegotiation becomes possible when the first input’s level reveals information about the state. (Actually, under some circumstances renegotiation might be possible even if only labor is monitored but capital is set first. See the end of the section for a discussion.) We initially restrict ourselves to only two cases, full revelation and no revelation

---

7 As modelled, only the owner makes renegotiation proposals. As Maskin and Tirole (1992) show, however, any contract that is renegotiation-proof in our setting is also renegotiation-proof in the game in which the manager makes proposals.
of information, but later we explain why such a restriction imposes no loss of
generality for our results. Full revelation occurs when \( k_1 \) is different from \( k_2 \) (the
separating case). Here, the choice of capital reveals the exact value of \( \theta \). Once \( \theta \) is
revealed, incentive constraints no longer bind and so renegotiation ensures that \( l \)
will be *ex post* efficient given \( \theta \) and \( k \). By contrast, no information is revealed when
\( k_1 \) equals \( k_2 \) (pooling), and so \( l \) can be more freely specified in the contract. In the
separating case, the optimal contract solves

\[
\text{max } \sum_{i=1}^{2} p_i V \left( \theta_2 f(k_i, l_i) - r_k - w_l + I_l \right)
\]

such that

\[
\sum_{i=1}^{2} p_i I_l \leq 0,
\]

\[
\theta_2 f(k_2, l_2) - r_k - w_l + I_2 \geq \theta_2 f(k_1, l_1) - r_k - w_l + I_1,
\]

and

\[
\left\{ \theta_1 f_2(k_1, l_1) \leq w \right\} \quad \left\{ l_1 \geq 0 \right\}
\]

where the inequalities in (4b) are complementarily slack. Inequalities (2b) and (3b) are,
by analogy with the earlier analysis, the individual-rationality and incentive constraints.
Formula (4b) can be thought of as a “renegotiation-proofness” constraint. It reflects
the fact that once the manager has set capital, (3b) no longer applies and so both the
manager and owner gain from setting \( l_1 \) efficiently.

In comparison to (1)–(4), this scheme forces the manager who claims \( \theta = \theta_1 \) to
choose \( k = k_1 \), making the cost of cheating higher (and also promoting better risk
sharing). However, \( l_1 \) is now constrained to be *ex post* efficient (via (4b)).

When capital is pooled, the optimal contract solves

\[
\text{max } \sum_{i=1}^{2} p_i V \left( \theta_2 f(k_i, l_i) - r_k - w_l + I_l \right)
\]

such that

\[
\sum_{i=1}^{2} p_i I_l \leq 0,
\]

\[
\theta_2 f(k_2, l_2) - r_k - w_l + I_2 \geq \theta_2 f(k_1, l_1) - r_k - w_l + I_1,
\]

and

\[
k_1 = k_2.
\]

Here, there is no renegotiation-proofness constraint because the owner learns nothing
from the choice of capital. Constraint (4c) is simply the requirement that capital be
pooled.
When can a one-variable scheme dominate an optimal two-variable scheme (without commitment against renegotiation)? The answer is “never” when this variable is capital, the first input to be set. Indeed, the solution to (1a)–(4a) is dominated by that of (1b)–(4b) (note that although we are not claiming the latter program is optimal, if it dominates the former, then a fortiori, so does the optimal two-variable scheme), because (4a) and (4b) both require ex post efficiency of labor, but (3b) imposes a greater cost on a cheating manager than (3a): in order to claim \( \theta = \theta_1 \), the \( \theta_2 \)-manager in the two-variable scheme, unlike in the one-variable case, must set capital below the ex post efficient level conditional on \( l_1 \) and \( \theta_2 \).

This argument, together with Proposition 1, enables us to draw the following simple conclusion:

**Proposition 2.** Under the hypotheses of Proposition 1, the optimal two-variable contract dominates any one-variable contract if \( w' < r \).

**Proof.** If \( w' \geq r \), then, given our hypotheses, the proof of Proposition 1 establishes that the parties are better off if capital rather than labor is monitored. But, as just explained, it is immediate that a two-variable contract is preferable to a one-variable contract where only capital is monitored. \( Q.E.D. \)

The more interesting comparison involves the labor scheme, that is, the solution to (1)–(4). To investigate when this scheme dominates two-variable contracts without commitment, we shall give conditions under which it is equivalent to the optimal two-variable scheme where parties can commit themselves not to renegotiate. The latter is the solution to (1b)–(3b), which we shall call the second-best optimum. Let \( (l_{**}^*, k_{**}^*, l_{**}^*) \) be the second-best optimal contract. In this contract, the manager’s expected payoff is

\[
 p_1 V\left( S_1 + p_2 (S_2 - S_{21}) \right) + p_2 V\left( (S_{21} + p_2 (S_2 - S_{21})) \right),
\]

\( (***) \)

where \( S_i = \theta_i f(k_i, l) - r k_i - w l_i, S_{21} = \theta_2 f(k_1, l_1) - r k_1 - w l_1, \) and all variables are evaluated at \( (l_{**}^*, k_{**}^*) \).

**Proposition 3.** Assume that \( \lim_{k \to \infty} f_1(k, l) = \lim_{k \to \infty} f_2(k, l) = \infty \). For any given \( \theta_i \), there exist positive constants \( a \) and \( b \) such that whenever \( \theta_1 > a \) and \( \theta_2 r < b \), there exists \( c > 0 \) such that for \( w < c \), the labor scheme strictly dominates any two-variable scheme (without commitment) and, in fact, attains the second-best optimum.

**Proof.** See the Appendix.

The idea behind Proposition 3 is that when the cost of capital is high and there is sufficient uncertainty about \( \theta \), it is (second-best) optimal to set both capital and labor at low levels in bad states for the purpose of risk sharing. Such an outcome can be approximated when only labor is monitored, moreover, since the high cost of capital will induce the manager to choose a low capital level anyway. However, when in addition the wage rate is low, the second-best optimum cannot be approximated if both factors are monitored and renegotiation is unpreventable. As we have observed, the constraint of renegotiation implies that capital must be pooled (i.e., set at the same level whether \( \theta \) is high or low) or else labor must be set ex post efficiently. Now, the former is inconsistent with second-best optimality: When \( \theta \) is low, capital is set at a low level in the second-best optimum (as we have noted), and when \( \theta \) is high, labor is employed at a high level, implying (since the marginal
product of capital rises with labor) that capital is too. But the latter is also inconsistent: The low wage implies that for ex post efficiency, employment must be high even when \( \theta \) is low, and this is in conflict with the employment distortion that second-best optimality demands.

From this discussion, we can conclude that allowing for two-variable contracts with partial pooling (e.g., where the \( \theta_2 \)-firm randomizes between \( k_2 \) and \( k_1 \)) would not affect the conclusion of Proposition 3. This follows from our demonstration that not only can a one-variable scheme dominate a two-variable contract, it can at the same time implement the second-best optimum (for an analogous result in a model of moral hazard, see Segal and Tadelis (1994)). Because this optimum entails no pooling at all, it necessarily dominates any two-variable scheme with partial pooling. We should add, however, that although the second-best optimality of the one-variable scheme is a feature that considerably simplifies our arguments, it is a much stronger property than needed to demonstrate the superiority of one-variable over two-variable contracts. Indeed, from continuity, it is clear that one-variable contracts will continue to dominate even when \( f^* \) is strictly positive, i.e., for parameter values for which they cannot attain the second-best optimum.

We have been implicitly assuming that if capital is not monitored, then there is no possibility of renegotiation. But this requires some qualification. If the owner knows for certain that the manager has already chosen capital (even though she cannot observe the capital level), then she may be able to induce a mutually beneficial renegotiation before labor is set: Because the right-hand side of (3), when \( i = 2 \), becomes \( \theta_2 f(k_2, l_1) - r k_2 - w_1 l_1 + I \), after the manager has made his capital decision, the "new" constraint (3) holds strictly (provided the original one was satisfied). Hence, the owner can propose raising \( l_1 \) and lowering \( I \), which improves efficiency. Thus, by ruling out renegotiation, we are implicitly assuming that the owner cannot determine whether or not the capital level has been set without actually monitoring capital. This assumption is very much in the spirit of the principle that one cannot keep track of a variable without setting up monitoring arrangements in advance (see the Introduction). To give a specific example, suppose that the technology requires capital to be set at date \( \tilde{i} \), where \( \tilde{i} \) is a random variable with an atomless distribution on the interval \([t_0, t_1]\). The realization of \( \tilde{i} \) is assumed to be private information to the manager. Suppose, moreover, that once capital is set, labor must be employed soon thereafter. Then at any given date, the probability that the manager has set capital but not labor is small, so renegotiation is prevented.

4. Monitoring cost saving

In this section we drop the assumption that labor is observable and assume instead that cost saving can be monitored directly (as in the Laffont and Tirole (1986) regulation model). The question to be answered is analogous to that in Section 3: Given that cost saving is monitored, when is it better to refrain from monitoring capital?

By analogy with the cheap factor principle, we might first ask which of capital or cost saving it is better to monitor in a single-variable scheme. This question was considered byMaskin and Riley (1985), who showed that cost-saving schemes (in their framework, "output" schemes) dominate capital schemes (in their framework, "input" schemes) whenever, in the first-best (full information) contract, greater cost-saving is always associated with higher input use. Of course, this hypothesis about the first-best

\[ \lim_{k \to 0} f(k, l) = \infty \] (see the statement of Proposition 3), \( k^* \) will, in fact, be big.
contract is not sufficient to make capital’s unobservability optimal because, when capital can be monitored, an additional constraint is imposed on the manager who falsely claims that $\theta = \theta_i$.

Specifically, when cost-saving alone can be monitored, the optimal contract solves

$$\max \sum_{q_i, k_i, l_i} p_i V(q_i - r k_i - w l_i + l_i)$$

such that

$$\sum_{i=1}^{2} p_i l_i \leq 0,$$

$$q_2 - r k_2 - w l_2 + l_2 \geq q_1 - r k_21 - w l_21 + l_1,$$

$$q_1 = \theta_2 f(k_21, l_21), \quad q_i = \theta_2 f(k_i, l_i), \quad i = 1, 2,$$

and

$$\frac{r}{w} = \frac{f_1(k_i, l_i)}{f_2(k_i, l_i)} = \frac{f_1(k_21, l_21)}{f_2(k_21, l_21)}, \quad i = 1, 2,$$

where $q_i$ is cost saving in state $\theta_i$. Here, (11) is the individual-rationality constraint, (12) is the incentive constraint, and the first equation in (13) is the requirement that if, in state $\theta_2$, the manager claims $\theta = \theta_i$, his choices of capital and labor be consistent with cost-saving $q_i$. The equations in (14) ensure that these choices be efficient relative to output.

When capital and cost saving can both be monitored, the optimal contract\(^9\) satisfies

$$\max \sum_{q_i, k_i, l_i} p_i V(q_i - r k_i - w l_i + l_i),$$

$$\sum_{i=1}^{2} p_i l_i \leq 0,$$

$$q_2 - r k_2 - w l_2 + l_2 \geq q_1 - r k_21 - w l_21 + l_1,$$

$$q_1 = \theta_2 f(k_21, l_21), \quad q_i = \theta_2 f(k_i, l_i), \quad i = 1, 2,$$

and

$$k_1 = k_2 \quad \text{or} \quad w = \theta_2 f_2(k_i, l_i), \quad i = 1, 2.$$\(^{14a}\)

Formula (14a) is the requirement of renegotiation-proofness.

Notice that when cost saving alone is monitored, the manager who falsely claims $\theta = \theta_i$ is free to choose his inputs to minimize costs. When capital is also monitored, he has to set $k = k_1$. However, for cost saving to be adjusted to improve risk sharing requires that $k_1 = k_2$; if $k_1 \neq k_2$, cost saving must be set ex post efficiently given $k_i$ and $\theta_i$.

\(^9\) Again, we restrict attention to separating and pure pooling contracts.
Although the contrast between monitoring capital or cost saving works somewhat differently from the earlier comparison of capital and labor, there is a clear counterpart to Proposition 3:

**Proposition 4.** Assume that \( \lim_{l \to \infty} f_1(k, l) = \lim_{k \to \infty} f_2(k, l) = \infty \). If \( V \) satisfies MURC, then, for given \( \theta_1 \), there exist positive constants \( a \) and \( b \) such that if \( \theta_2/\theta_1 > a \) and \( \theta_2/r < b \), there exists \( c > 0 \) such that for all \( w < c \), the optimal cost-saving scheme implements the second-best optimum (i.e., the solution to (10a)–(13a)) and strictly dominates any two-variable contract without commitment against renegotiation.

**Proof.** Let \( \langle k^{**}, q^{**}, l^{**} \rangle \) and \( \langle k^0, k^0_2, q^0, l^0 \rangle \) be the solutions to (10a)–(13a) and (10)–(14) respectively. Paralleling the argument in the proof of Proposition 3, we can show that there exist \( a \) and \( b \) such that, for \( \theta_2 > a \) and \( \theta_2/r < b \), \( k^{**} = k^0 = k^0_2 = 0, k^* = k^*_2 > 0 \), and \( q^{**} = q^0 = 0 \). By contrast, there exists \( c \) small enough such that for \( w < c \), the solution to (10a)–(14a) \( \langle k^*, q^*_2, l^* \rangle \) cannot obtain the second-best optimum. If capital is pooled \( k^* = k^*_2 \), we obtain a contradiction with the fact that \( k^{**} \neq k^*_2 \). If capital is not pooled, then for \( \text{ex post} \) efficiency, \( q^* \) must be set at a positive level because \( w \) is low, which is in conflict with the above choice of \( q^{**} \).

Q.E.D.

Proposition 4 establishes that when \( r \) is big enough, capital and output are set at inefficiently low levels in state \( \theta_1 \) in the second-best solution, because the allocative loss of doing so is outweighed by the gain that derives from the relaxation of the incentive constraint. A cost-saving scheme can replicate this outcome: When output is set low in state \( \theta = \theta_1 \), the manager finds it optimal to keep the level of capital low. However, the two-variable scheme cannot attain the second-best when there is sufficient uncertainty (i.e., \( \theta_2 \) is big enough relative to \( \theta_1 \)) and a sufficient difference in factor costs: Capital pooling imposes too great an allocative loss, whereas full separation of capital gives rise to renegotiation that interferes with risk sharing.

5. **Concluding remarks**

- In this article we have studied the optimal set of screening variables in adverse selection problems with renegotiation. Specifically, we have considered when it is optimal to restrict the observability of inputs to a cost-saving process.

  Of course, our model is very simple, but we believe the principles involved are quite general. In particular, there should be little difficulty in finding suitable extensions of the results for more than two states of nature and more than two screening variables.

  A particularly interesting generalization might be to a model with a longer time horizon. Intuitively, the longer a contract lasts, the more troublesome the constraints imposed by renegotiation. This suggests that contracts should, if anything, become simpler (i.e., less contingent) as their duration grows. It would be worthwhile to confirm this conjecture.

  Another potentially interesting extension would be to a setting in which more than one screening variable is determined at the same time (e.g., several different kinds of capital are set more-or-less simultaneously.) In this case, renegotiation would constrain pooling to be coordinated across all these variables. Because such coordination could be costly, the case for restricting observability might, once again, be strengthened.

  Our analysis suggests that parties may gain from restricting \textit{ex ante} what the uninformed party can observe, even when monitoring is costless.\(^{10}\) Hence, our explanation for relatively simple contracts contrasts with those based on exogenous costs of

\(^{10}\) By contrast, in the auditing literature (e.g., Baron and Besanko, 1984), one tries to explain when \textit{costly} auditing is used in equilibrium, as a way to improve the tradeoff between allocative efficiency on the one hand and rent extraction or insurance on the other.
complexity or exogenous limits on verifiability by third parties, since here renegotiation is the only constraint faced by the contracting parties.\footnote{11} Our story also differs from the idea that complex renegotiation-proof contracts can be replicated by simple contracts that are then renegotiated in equilibrium (see Huberman and Kahn (1988) and Aghion, Dewatripont, and Rey (1994)).

Appendix

- Proofs of Propositions 1 and 3 follow.

Proof of Proposition 1. Define $S_i = \theta_i f(k_i, l_i) - r_k - w_l$, $S_{2ii} = \theta_i f(k_{2i}, l_{2i}) - r_{k2i} - w_{l2i}$, and

$$S_{2ik} = \theta_i f(k_{2i}, l_{2i}) - r_k - w_{l2i}.$$ 

Then, since (2) and (3) ($i = 2$) are binding at the optimum, the manager’s program under labor monitoring can be rewritten as

$$\max p_1 V(S_i + p_2(S_2 - S_{2ii})) + p_2 V(S_{2ii} + p_2(S_2 - S_{2ii})) \quad \text{subject to (4).} \quad (A1)$$

And, similarly, the manager’s problem under capital monitoring consists of:

$$\max p_1 V(S_i + p_2(S_2 - S_{2ii})) + p_2 V(S_{2ik} + p_2(S_2 - S_{2ii})) \quad \text{subject to (4a).} \quad (A2)$$

Define $G(\lambda)$ on $[0, 1]$ so that

$$G(\lambda) = p_1 V\left(\lambda(S_{2ii} + p_2(S_2 - S_{2ii})) + (1 - \lambda)(S_0 + p_2(S_2^0 - S_{2ii}^0))\right) + p_2 V\left(\lambda(S_{2ii} + p_2(S_2 - S_{2ii})) + (1 - \lambda)(S_{2ik} + p_2(S_2 - S_{2ik}))\right),$$

where expressions superscripted by 0 are evaluated at $(l_0, k_0, k_0^0)$ and those superscripted by 00 are evaluated at $(l_0^0, k_0^0, l_0^0)$. By the mean value theorem, the difference between (A1) and (A2), $G(1) - G(0)$, can be expressed as

$$G'(\lambda^*) = p_1 V'(1)[S_0 - S_{00}] + p_2[p_2 V'(1) - V'(2)][S_2^{10} - S_{20}]$$

for some $\lambda^* \in [0, 1]$, where $V'(1)$ and $V'(2)$ are the derivatives of $V$ evaluated, respectively, at $\lambda(S_{2ii} + p_2(S_2 - S_{2ii})) + (1 - \lambda)(S_0 + p_2(S_2^0 - S_{2ii}^0))$ and $\lambda(S_{2ii} + p_2(S_2 - S_{2ii})) + (1 - \lambda)(S_{2ik} + p_2(S_2 - S_{2ik}))$. (Note that $S_0 = S_{00}$.) Fix $(r^*, w^*)$, where $r^* > w^*$. We shall show that for $\theta_1$ big enough relative to $\theta_2$, (A3) is positive for $r = r^*$ and all $w \in [w^*, r^*]$. By symmetry, $G(1) = G(0)$, i.e., (A3) is zero when $w = r^*$. Next, observe that as $\theta_1$ tends to infinity, $S_{2ii}^0 < S_{2ii}^0$ and $S_{2ii}^0 - S_{2ii}^i$ converge to infinity uniformly for all $w \in [w^*, r^*]$. $(S_0^0)$ and $S_0$ are bounded from above by $\max(\theta_i f(k, l) - r_k - w_l)$. Because $f_i(0, 0) > 0$, however, $S_{2ii}^0$ goes to infinity as $\theta_2 \to \infty$. Similarly, so does $S_{20}$. Hence the difference between the arguments of $V'(1)$ and $V'(2)$ goes to infinity uniformly, so if we let $x$ be the argument of $V'(1)$ and $x + y$ be the argument of $V'(2)$, we conclude that $y \to \infty$ uniformly. Whether or not $x \to \infty$, MURC then implies that for $\theta_1$ big enough, $V'(2)/V'(1) < \frac{1}{2}$, so $V'(1) - V'(2) < \frac{1}{2}$ for all $w \in [w^*, r^*]$. Now, if $\theta_1$ is big enough, then $l_0^0 = k_0^0 = 0$ for all $w \in [w^*, r^*]$. (For $\theta_2$ big enough, $V'(2)/V'(1)$ is nearly zero, so the first-order condition for $l_0^0$ is nearly $\theta_2 f(k_0^0, l_0^0) - w_1 - p_2(\theta_2 f(k_0^0, l_0^0) - w) \leq 0$, with equality if $l_0^0 > 0$. But when $\theta_2$ is big enough so that $\theta_1 < p_2 \theta_2$, the left-hand side of this inequality is necessarily negative—since $k_0^0 \leq k_0^0$—and so $l_0^0 = 0$. The argument for $k_0^0$ is similar.) Hence, for such $\theta_2$,

$$\frac{\partial}{\partial w}(S_i - S_0^0) = l_0^0$$

and

\footnote{11} Spier (1992) has argued that another reason why contracts may be left less than fully contingent is that for a party to propose a particular contingency clause may signal information it would prefer to keep private.
As \( \theta_2 \to \infty \), \( l_{21}^\infty \) tends to infinity (since \( f_2 > 0 \)). Hence, for \( \theta_2 \) big enough, \( l_{10}^\infty /l_{21}^\infty < \frac{1}{2}p_2 \) for all \( w \in [w^*, r^*]. \)

We conclude that (A3) is positive for all \( w \in [w^*, r^*]. \) \( \Box \).

Proof of Proposition 3. We first claim that there exist \( a \) and \( b \) such that if \( \theta_2 > a \) and \( \theta_2/r < b \), \( k_1^\infty = 0. \) Suppose not. Then there exists a sequence \( \{ (\theta_2^*, r^*) \} \) such that \( \theta_2^* \to \infty, \theta_2/r^* \to 0 \), and for which the corresponding solution \( \{ (k_1^*, l_1^*) \} \) to (1b)-(3b) satisfies \( k_1^* > 0 \) for all \( n \). Assume first that there exists \( l > 0 \) and a subsequence \( \{ l_1^* \} \) of \( \{ l_1^* \} \) such that \( l_1^* \geq l \) for all \( n \). Now along the subsequence, \( S_{21} - S_{11} \geq (\theta_2^* - \theta_2)f_0(0, l) \).

Hence, for any \( \epsilon > 0 \), MURC implies that there exists \( n \) big enough so that \( V'(2)/V'(1) < \epsilon \), where \( V(1) \) and \( V(2) \) are the Vs in the first and second terms, respectively, of (**). Differentiating (**) with respect to \( k_1 \), we obtain the first-order condition

\[
0 > v = \frac{\partial}{\partial w} (S_{21}^{\theta_2} - S_{11}^{\theta_2}) = -l_{21}^{\theta_2}.
\]

We conclude that (A3) is positive for all \( w \in [w^*, r^*]. \) \( \Box \).

Proof of Proposition 3. We first claim that there exist \( a \) and \( b \) such that if \( \theta_2 > a \) and \( \theta_2/r < b \), \( k_1^\infty = 0. \) Suppose not. Then there exists a sequence \( \{ (\theta_2^*, r^*) \} \) such that \( \theta_2^* \to \infty, \theta_2/r^* \to 0 \), and for which the corresponding solution \( \{ (k_1^*, l_1^*) \} \) to (1b)-(3b) satisfies \( k_1^* > 0 \) for all \( n \). Assume first that there exists \( l > 0 \) and a subsequence \( \{ l_1^* \} \) of \( \{ l_1^* \} \) such that \( l_1^* \geq l \) for all \( n \). Now along the subsequence, \( S_{21} - S_{11} \geq (\theta_2^* - \theta_2)f_0(0, l) \).

Hence, for any \( \epsilon > 0 \), MURC implies that there exists \( n \) big enough so that \( V'(2)/V'(1) < \epsilon \), where \( V(1) \) and \( V(2) \) are the Vs in the first and second terms, respectively, of (**). Differentiating (**) with respect to \( k_1 \), we obtain the first-order condition

\[
p_{1}V'(1)[(\theta_1, \theta_2)f_0(k_1^*, l_1^*) - p_1r] + p_{1}p_{2}V'(2)[\theta_2f_0(k_1^*, l_1^*) - r] \leq 0, \tag{A4}
\]

where strict inequality obtains only if \( k_1^*= 0. \) But for \( n \) big enough along the subsequence,

\[
p_{1}(\theta_1, \theta_2) + p_1p_2 \epsilon \theta_2 < 0
\]

(provided that \( \epsilon < 1 \)). Hence, for such \( n \), the left-hand side of (A4) is negative, so \( k_1^* = 0 \), a contradiction of how the sequence \( \{ (\theta_2^*, r^*) \} \) was chosen. Hence, \( l_1^* \to 0 \), and so for large \( n \) the left-hand side of (A4) is arbitrarily close to

\[
p_{1}V'(1)[\theta_1f_0(l_1^*, 0) - w] - p_{1}p_{2}V'(2)[\theta_2f_0(0, l_1^*) - r].
\]

Moreover, \( V'(1) \geq V'(2) \); so, since \( f_1(0, 0) < \infty, k_1^* = 0 \) for \( n \) big enough, a contradiction. We conclude that for \( a \) big enough and \( b \) small enough, \( k_1^\infty = 0 \) after all. In that case, \( l_1^\infty \) is determined by the following first-order condition:

\[
p_{1}V'(1)[\theta_1f_0(0, l_1^\infty) - w] - p_{1}p_{2}V'(2)[\theta_2f_0(0, l_1^\infty) - r] \leq 0. \tag{A5}
\]

Turning to the labor-monitoring scheme, recall from (4) that \( \theta_2f_0(k_2^0, l_1^0) = r \) and \( \theta_1f_0(k_1^0, l_1^0) = r \).

Differentiating (A1) with respect to \( l_1 \) and evaluating at \( l_1 = l_1^0 \), we obtain the first-order condition

\[
p_{1}V'(1)[\theta_1f_0(k_1^0, l_1^0) - w] - p_{1}p_{2}V'(2)[\theta_2f_0(k_2^0, l_1^0) - r] \leq 0,
\]

where primes denote derivatives with respect to \( l_1 \), (and \( k_2^0, k_1^0 \) depend on \( l_1 \) through the first-order conditions (4)). If \( k_1^0 \neq 0 \), then \( k_2^0 > 0 \) and so \( \theta_2f_0(k_2^0, l_1^0) = r \). Hence \( \theta_2f_0(k_2^0, l_1^0) - r)k_1^0 = 0 \) and similarly

\[
\theta_1f_0(k_1^0, l_1^0) - r)k_1^0 = 0.
\]

Thus, the above inequality reduces to

\[
p_{1}V'(1)[\theta_1f_0(k_1^0, l_1^0) - w - p_2(\theta_2f_0(k_2^0, l_1^0) - w)] + p_{1}p_{2}V'(2)[\theta_2f_0(k_2^0, l_1^0) - w] \leq 0. \tag{A6}
\]

We claim that there exist \( a \) and \( b \) such that if \( \theta_2 > a \) and \( \theta_2/r < b \), \( k_2^0 = 0. \) Suppose to the contrary that there exists a sequence \( \{ (\theta_2^*, r^*) \} \) as in the first paragraph such that the corresponding solutions

\[
\{ (k_1^*, k_2^0, l_1^*) \}_{i=1,2}
\]

satisfy \( k_2^0 > 0 \) for all \( n \). If there exists \( l > 0 \) such that, for some subsequence \( \{ l_1^i \}, l_1^i \geq l \) for all \( n \), then, as above, \( V'(2)/V'(1) \) becomes arbitrarily small for \( n \) big enough; so, from (A6), \( l_1^0 = 0 \) for \( n \) big enough, a contradiction. Hence \( l_1^0 \to 0 \) after all. But then, from the first-order conditions (4) for \( k_2^0 \) and \( k_1^0 \), we have \( k_1^0 = 0 \) for \( n \) big enough, a contradiction. We conclude that the claim is true. Thus for \( (\theta_2, r) \) satisfying the inequalities, \( k_2^0 = 0 \) and so \( k_2^0 = 0. \) Formula (A6) then becomes

\[
p_{1}V'(1)[\theta_1f_0(0, l_1^0) - w - p_2(\theta_2f_0(0, l_1^0) - w)] + p_{1}p_{2}V'(2)[\theta_2f_0(0, l_1^0) - w] \leq 0.
\]

But this is the same as (A5). Hence, we can find \( a, b \) such that if \( \theta_2 > a \) and \( \theta_2/r < b \), then
$k^*_1 = k^*_2 = k^{**} = 0$ and $l^*_0 = l^{**}$.

What about two-variable schemes without commitment? If, given $a$ and $b$ as above, then for $(b_r, r)$ satisfying $b_r > a$ and $b/r < b$, $k^{**} > 0$ and $k^{**} = 0$, so that if the two-variable scheme entails pooling (i.e., it solves (1c)-(4c)), it cannot be second-best optimal. If the two-variable scheme entails complete separation, it solves (1b)-(4b). Let $(k^*_1, l^*_0)_{r=1,2}$ be the solution. Then, for $c < \theta_f(0, 0)$ and $w < c$, $l^*_0$ satisfies $\theta_f(0, l^*_0) = w$. Now if $l^{**} = 0$, then we are done, since clearly $l^*_0 \neq l^{**}$. Thus suppose that $l^{**} > 0$. In this case, $V'(1) > V'(2)$ in (AS), the condition that determines $l^{**}$. Hence, from (AS), $l^{**}$ does not satisfy $\theta_f(0, l^{**}) = w$, and we conclude again that $l^*_0 \neq l^{**}$. Q.E.D.

References


