Incentives and Standards in Agency Contracts

Robert G. Chambers[1] and John Quiggin[2]

December 22, 2000

[1]Professor, Department of Agricultural and Resource Economics, University of Maryland, College Park and Adjunct Professor, Faculty of Agriculture, University of Western Australia.

[2]Australian Research Council Fellow, Australian National University

Abstract

This paper studies the structure of state-contingent contracts in the presence of moral hazard and multi-tasking. Necessary and sufficient conditions for the presence of multi-tasking to lead to fixed payments instead of incentive schemes are identified. It is shown that the primary determinant of whether multi-tasking leads to higher or lower powered incentives is the role that noncontractible outputs play in helping the agent deal with the production risk associated with the observable and contractible outputs. When the noncontractible outputs are socially undesirable and risk substitues, standards are never optimal. If the noncontractible outputs are socially desirable, standards are never optimal if the noncontractible outputs play a risk-complementary role.

Key words: incentives, multi-tasking, agency, risk complementarity, risk substituability
D-82, L-23, L-50

The central tenet of contract theory is that agents respond to incentives in a self-interested manner. A natural consequence of this assumption is that a principal who rewards agents should be prepared for the agents to undertake actions in pursuit of those rewards that may be counter to his own objectives. Hence, the literature on compensation systems has long recognized that inappropriately designed incentive schemes can lead to counterproductive (from the principal's perspective) actions by the agent (Lawler 1971; Kerr 1975). Similarly, economic regulation of firms and industries often has unintended and unforeseen consequences.

The notion that basing agent remuneration on objective performance measures can be harmful to the principal's interest in other areas where objective performance measures are unavailable is central to the multi-task moral hazard and performance measurement literatures (Holmström and Milgrom 1991; Baker 1992). Some examples illustrate. Holmström and Milgrom (1991) cite the controversy surrounding linking teacher pay to student performance on standardized tests. Teacher 'accountability' schemes have been introduced in a number of U.S. states. Advocates argue that it introduces incentives for increased teacher effort. Opponents counter that it encourages teachers to neglect educational activities that are not objectively measurable in favor of activities designed solely to raise student scores on standardized tests. While debate continues on this issue, little doubt exists that teachers respond to these incentives,[1] and there is some limited evidence that these schemes do affect student performance (Ladd, 1999).

Another example emerges from the common observation that manufacturing firms rather infrequently employ piece-rate systems. One explanation is that piece-rate rewards can provide workers adverse incentives for maintaining the quality of the firm's machinery (Alchian and Demsetz, 1972; Prendergast, 1999).

Chambers and Quiggin (1996) suggest an example from a regulatory framework. Both the European Community and the United States have long provided subsidies and price support to farmers. Such support schemes are increasingly recognized to conflict with environmental objectives. For example, providing farmers with formal or informal crop insurance can encourage riskier production activities which also degrade the environment in the form of chemical runoff and soil erosion.

In each of these examples, an incentive scheme is based on an observable and objective

[1] Two examples taken from recent newspaper stories illustrate. In June 2000, the Washington Post reported two separate incidents involving allegations that teachers and school principals had been involved in assisting students to cheat on standardized examinations (Schulte, 2000; Eggen, 2000). On a more positive note, the New York Times reports on the success of a cash incentive scheme for teachers in North Carolina (Steinberg, 2000).

performance criterion, even though the principal also cares about noncontracted aspects of the agent's performance. Starting with the seminal papers of Holmström and Milgrom (1991) and Baker (1992), theorists have increasingly speculated that the existence of multi-task concerns may partially explain the tendency of employment contracts to specify fixed wages, or, more generally, the commonly noted presence of muted incentives within firms. In a regulatory context, one might similarly speculate that the frequent tendency toward command and control regulation in place of taxes or subsidies may emerge from multi-task concerns.[2]

This paper revisits the issue of multi-task agency problems, or multi-tasking for short, from a state-contingent perspective. It was long thought that the state-space approach to principal-agent problems was intractable (Hart and Holmström, 1988). As a consequence, much of the theoretical discussion of moral hazard has been cast in the 'parametrized distribution formulation' (Hart and Holmström, 1988). However, Quiggin and Chambers (1998) and Chambers and Quiggin (2000) have recently argued that this seeming intractability emerged from an implausible specification of state-contingent technologies in early principal-agent treatments, and that a more plausible specification leads to a tractable and easily manipulable moral-hazard model. This paper extends that observation to the study of multi-tasking.

When viewed from the perspective of a plausible state-contingent production technology, the issue of multi-tasking takes a different flavor. Characteristics of the stochastic technology assume center stage. The effect of multi-tasking on contract design in a state-contingent framework hinges on the role that the noncontractible activities play in allowing the agent to control the production risk associated with the contractible outputs. The emphasis, therefore, switches from agent 'tasks' to the outcomes from these tasks. The crop-runoff example illustrates: the tasks might include fertilization of the crop, weeding, and purely abating activities. The outputs would include the realized crop and the actual run-off pollution. Presumably, a regulator would be more interested in the run-off than in the activities undertaken to control that run-off. Our focus is on those cases where it is more reasonable to presume that the principal is primarily concerned with outcomes from the tasks rather than with the tasks themselves.

In what follows, we first present a graphical overview of our basic ideas. From there, we turn to a formal specification of the model. The model follows our earlier work on the state-contingent representation of moral-hazard as detailed in Quiggin and Chambers (1998) and Chambers and Quiggin (2000). The basic analytic device is to recognize that the state-space

[2] Slade (1996) examines some of the empirical implications of multi-task theory. Prendergast (1999) provides a convenient summary of the theoretical and empirical work on agent response to incentives.

moral hazard problem can be informatively rewritten as a problem of hidden knowledge about the state of Nature that occurs.

After the model is speci…ed and some preliminary results are developed, we …rst examine when …xed payments are optimal in the presence of multi-task considerations. Holmström and Milgrom (1991) appear to have been the …rst to have identi…ed conditions under which …xed payments can be optimal in the presence of moral hazard. They show that, when agent e¤ort in various tasks are perfect substitutes, …xed wages can dominate incentive schemes. The intuition is relatively simple. If tasks are perfect substitutes, introducing incentives for a subset of the tasks leads the agent to direct all his attention to the rewarded tasks. If the principal values some of the nonrewarded tasks, he may be better o¤ not using performance indexed incentive schemes.

Our focus is somewhat di¤erent. For the sake of concreteness, we presume that the noncontractible outputs are socially damaging, e.g., pollution, teacher collusion on student cheating, damage to the …rm's capital stock, consumer ill-will. Our result is that, provided an incentive scheme is implementable, a …xed payment is only optimal for strictly monotonic agent cost structures when the noncontractible output plays a 'risk-complementary' role in the production of the contractible output. By risk-complementary, we essentially mean that higher levels of the noncontractible output are associated with riskier outcomes with respect to the contractible output. (This is made precise below.) For weakly monotonic cost structures, …xed payments are the norm if the production of the contractible outputs is technically e¢cient. If incentive schemes are to be used with weakly monotonic cost structures, they must also encompass technical ine¢ciency.

We then take up the issue of higher- vs. lower-powered incentive schemes. Here our method is to examine how the introduction of multi-task concerns a¤ects the contracts that would have been chosen in the absence of multi-tasking. We show that higher powered incentives tend to emerge when the noncontractible outputs play a risk-substituting role in the production of the contractible outputs. Lower powered incentives emerge when the noncontractible outputs play a risk-complementary role.

After a general discussion of higher- vs. lower-powered incentive schemes, we then apply our ideas to the crop-runo¤ problem. We deduce a number of speci…c results for that case, and then the paper closes.

1 An Overview

Consider Figure 1. There we present a representative isocost curve for the agent. Production and consumption in state of Nature 2 are measured against the vertical axis, while production and consumption in state of Nature 1 are measured along the horizontal axis. In the absence of moral hazard, the principal would ask the agent to produce the …rst-best state-contingent output mix, which equates the relative odds of the states occuring to the state-contingent marginal rate of transformation. This is illustrated by the ray labelled FB in the …gure, which passes through the point where the fair-odds line (not drawn) is tangent to the agent's isocost curve. As a reward, the principal o¤ers a non-stochastic payment, which in state-contingent terms is depicted as a point on the bisector. Thus, the agent is fully insured against income risk and maximizes the expected pro…t from production. State-contingent payments do not vary with state-contingent production.

If the principal faces single-task moral hazard, such a contract structure leads to agent shirking. The principal must tie payments to production to provide incentives for agent e¤ort. Relative to the …rst-best, the principal asks the producer to produce a less dispersed state-contingent output vector in return for a more dispersed payment vector (Quiggin and Chambers, 1998).[3] In Figure 1, this means the state-contingent production vector pivots to the right, say, to the ray depicted by A, while the state-contingent payment vector pivots to the left from the bisector, say, to the ray depicted by B.

How then do we expect the presence of multi-task moral hazard to a¤ect the incentive structure for the contractible task when the economic e¤ects of the other tasks are hard to measure or verify? Do we expect the principal to respond by muting the incentives he introduced in the single-task problem or by introducing higher powered incentives than in the single-task problem?

Viewed from the perspective of Quiggin and Chambers (1998), this is akin to asking whether or not multi-tasking creates countervailing incentives. In the adverse selection literature, if there exist countervailing incentives, one expects bunching or pooling (Lewis and Sappington 1989). In a moral hazard context, bunching or pooling translates into low powered or performance-independent incentive schemes. Whether such schemes emerge depends on how the multiple tasks interact in the agent's response to a risky production situation.

Consider Figure 1 in the context of the crop-runo¤, nonpoint source pollution problem discussed in the introduction. Suppose that when the producer undertakes a risky, but potentially high yielding, production strategy, he also emits large amounts of runo¤ pollution. An example here is given by nitrogen runo¤ associated with the heavy use of chemical fertilizer. Intuitively, one then expects that the principal (the regulator) would respond by asking the agent to produce a less risky state-contingent production vector than in the single-task case. Naturally, one expects the regulator to achieve this by muting the incentives for

[3]Strictly speaking, demonstrating this result requires assuming the state-contingent cost structure is homothetic.

producing output. If the pollution problem is severe enough, the principal might even be even respond by offering a fixed payment for a fixed output. In terms of pollution regulation, this means that the principal imposes a production standard in place of an incentive (tax or subsidy based) scheme. We show, in fact, that a production standard, or more generally a fixed payment, is only optimal if runoff pollution is associated with riskier production strategies.

On the other hand, suppose that larger amounts of runoff are associated with less risky production activities. For example, one can now think in terms of large-scale applications of chemical pesticides and their associated contamination of the groundwater. Pesticides, by their nature, do not raise maximal output so much as they control damage to potential output. Intuitively, one now expects the principal to respond by asking the agent to produce a more risky state-contingent output vector than in the single-task case. The principal could implement such a state-contingent output vector by introducing a payment scheme with higher-powered incentives than in the single-task case. Moreover, it seems implausible that a fixed payment or a production standard would ever be optimal in these circumstances. We show this formally below.

2 The Model

There are two individuals: a principal and an agent. The agent produces two outputs, a socially bad output, say pollution, and a valuable output under conditions of production uncertainty. Only the agent engages in productive activity. Uncertainty is modelled by 'Nature' making a choice from a set of two alternatives, $\Omega = \{1, 2\}$. Production relations are governed by a state-contingent input correspondence (Chambers and Quiggin, 2000), $X : \Re^2_+ \times \Re_+ \to \Re^n_+$, defined by

$$X(z, p) = \left\{ x \in \Re^n_+ : x \text{ can produce } (z, p) \right\}, \quad z \in \Re^2_+, p \in \Re_+.$$

Here x represents an input vector that is committed prior to the resolution of uncertainty, i.e., before Nature makes its choice from Ω, and z is a vector of state-contingent output also chosen before Nature makes its choice. In line with much of the recent literature on multi-tasking, one can alternatively interpret x as the various 'tasks' undertaken by the agent. If Nature picks state s then the ex post or realized value of the output is z_s. p is a non-stochastic output,[4] which is socially undesirable, associated with the production of the

[4] p is taken as non-stochastic to preserve notational simplicity. This specification, however, is consistent with the treatment in other multi-tasking studies where 'tasks' are typically viewed as non-stochastic activ-

state-contingent output vector and the input committal.[5]

The principal is risk-neutral and does not directly value the agent's effort, but he does value the outputs that emerge from that effort.[6] Damage associated with p is given by a twice differentiable, strictly increasing, and strictly convex function $m(p)$. The agent's preferences depend upon his payment from the principal, which we denote as y, the vector of inputs, x, that he commits, and the level of the socially undesirable output, p. His ex post utility function is given by

$$w(y; x) = u(y) - g(x; p)$$

where $u : \Re_{++} \to \Re$ is a strictly increasing, strictly concave, and twice differentiable function, and $g : \Re_{+}^{n+1} \to \Re$ is everywhere continuous, increasing in x and p, and strictly convex.[7] Both u and g satisfy the von Neumann-Morgenstern postulates. The principal and the agent share the same probabilities of a state s occurring and that probability is denoted by $\pi_s \in \Re_{++}$ $(s = 1; 2)$ with $\pi_1 + \pi_2 = 1$.

Ex post, the principal cannot observe either the state of nature or level of p. Hence, the informational asymmetry between the principal and the agent encompasses both hidden action (input committal, p emitted) and hidden knowledge (ex post about the state of nature that occurs) on the part of the agent. What is observable, and by assumption contractible to both parties, is the ex post output.

Two indirect representations of the state-contingent input correspondence prove useful. The first is the agent's effort-cost function for a given vector of state-contingent output and a pollution level. It is defined by

$$c(z; p) = \min_x \{g(x; p) : x \in X(z; p)\}$$

ities undertaken by the agent. More generally, however, our approach allows p to be stochastic at the cost of increased notational complexity.

[5] p is taken to be socially undesirable for the sake of clarity and because it reflects the role that p plays in our theoretical example. As pointed out below, our results can be extended in an obvious fashion to the case where p is socially desirable.

[6] This may seem to represent somewhat of a departure from much of the recent literature on multitasking where the principal is assumed to have direct preferences over some of the agent's tasks. However, the difference at best is only semantic and corresponds closely to the semantic difference between inputs, outputs, and netputs in the axiomatic literature on production.

[7] g is taken to be increasing in p to avoid the agent having pathological preferences over the socially undesirable output. If p were socially desirable, the parallel assumption would be that g is decreasing in p. However, it is easy to see that these assumptions could be easily relaxed if one's interest is in modelling purely 'anti-social' behavior.

if there exists an $x \in X(z;p)$ and ∞ otherwise. Hence, for fixed $(y;z;p)$, the agent's ex post welfare is

$$u(y) - c(z;p).$$

It is always assumed that the agent's production of $(z;p)$ results in jointness in $c(z;p)$. More precisely, $c(z;p)$ is not additively separable in z and p. Given the structure of the agent's evaluation of effort, this jointness can arise from several sources. It can be embedded in the production technology itself (formally, the production technology would not be input nonjoint). Alternatively, the input correspondence could be input nonjoint, but the nonlinearity of g could result in this jointness. It is important to recognize, however, that even in the presence of linear pricing of the inputs, x, one generally does not expect $c(z;p)$ to be additive.

The second cost function that we consider is the one that is relevant when the agent privately chooses the level of p associated with a given vector of state-contingent outputs. Denote the agent's private cost function by

$$C(z) = \mathrm{Min}_p \{c(z;p)\}.$$

We assume that $C(z)$ is strictly increasing and positively linearly homogeneous in z.[8] Let

$$p(z) = \arg\min_p \{c(z;p)\}.$$

Both C and $p(z)$ are assumed to be everywhere smoothly differentiable.

If $c(z;p)$ does not exhibit jointness, then the agent's choice of z and p are independent and

$$p(z) = p;$$

for all z. The principal, therefore, cannot use an output-based incentive scheme to implicitly regulate p. In effect, the principal is then reduced to delegating the authority for the optimal choice of p completely to the agent.[9]

Following Chambers and Quiggin (1997, 2000), we denote by $\preceq_{\pi}$ a partial ordering of $\Re^2_+$ that ranks vectors of state-contingent output vectors with the same mean. The notation

$$z \preceq_{\pi} z'$$

[8] Homogeneity is imposed to rule out the possibility of states of Nature cycling between 'good' and 'bad' (Quiggin and Chambers, 1998; Chambers and Quiggin, 2000, Chapter 4).

[9] In fact, if one were to use our model to study the optimal delegation of authority for certain tasks by the principal to the agent, the optimality of various degrees of delegation would depend critically upon the degree of jointness of $c(z;p)$.

means that $\sum_{s\in\Omega}\pi_s z_s = \sum_{s\in\Omega}\pi_s z_s^0$, and that z is less risky in the Rothschild-Stiglitz sense than z^0. $p(z)$ is risk-complementary at z if

$$z \preceq_\pi z^0 \Rightarrow p(z) \leq p(z^0)$$

and risk-substituting at z if

$$z \preceq_\pi z^0 \Rightarrow p(z) \geq p(z^0).$$

Intuitively our notion of risk-complementarity is that the producer responds to a local increase in the riskiness of the output vector by setting $p(z)$ at a higher level (Chambers and Quiggin, 2000, Chapter 4). Notice that $p(z)$ is risk-complementary at z only if

$$\left(\frac{p_1(z)}{\pi_1} - \frac{p_2(z)}{\pi_2}\right)(z_1 - z_2) \geq 0,$$

where subscripts on functions denote partial derivatives (Chambers and Quiggin, 2000, Chapter 4). $p(z)$ is risk-substituting only if the above inequality is reversed.[10]

Our production technology is general enough that either state 1 or state 2 could be the 'good' state of Nature in the sense that a risk-neutral individual facing this technology would choose to produce a higher output in that state of Nature. To order states of Nature, therefore, we assume without further loss of generality[11] that

$$\text{(1)} \qquad \frac{C_2(z, z)}{\pi_2} \leq \frac{C_1(z, z)}{\pi_1}, \quad z \in \Re_+.$$

Therefore, a risk-neutral individual facing this technology would always choose $z_2 \geq z_1$. Following, Quiggin and Chambers (1998), a state-contingent output vector is defined as monotonic if $z_2 \geq z_1$.

Suppose

$$\frac{p_1(z)}{\pi_1} - \frac{p_2(z)}{\pi_2} \leq 0, \quad \forall z.$$

Then p is risk-complementary for monotonic z, and risk-substituting for non-monotonic z. The converse also holds.

In the state-contingent framework, it is natural to think of different state-contingent outputs as having different relative input-intensities. For example, returning to the chemical-runoff problem, if the set of states of nature includes states favorable to insect infestations,

[10] To derive these results, take a differentiably small multiplicative spread of z under the assumption that p is risk complementary or risk substituting (Chambers and Quiggin, 2000, Chapter 4).

[11] This is purely a convention. By homogeneity, either (1) or its reversal must hold.

output in those states will be relatively pesticide-intensive. Whether pesticides are risk complements or risk substitutes depends on whether the relevant states of nature have relatively high or relatively low output. For any given technology, this, in turn, will depend on whether the output vector under consideration is monotonic or non-monotonic.

With the ability to monitor p and the ex post state of Nature, the principal chooses p and z to maximize

$$\pi_1 z_1 + \pi_2 z_2 - c(z;p) - m(p).$$

Optimality would require making a non-stochastic payment y to the agent such that

$$u(y) - c(z;p) = \bar{u}$$

where $\bar{u}$ represents the agent's reservation utility, while choosing z and p to maximize expected surplus. The solution to this problem will be referred to as the first-best.

If pollution is a risk substitute for monotonic z; then, since the socially optimal p will be lower than the privately optimal p, the first-best optimal z will be riskier than the private optimum, and hence will also be monotonic. Hence, a sufficient condition for the first best z to be monotonic is that pollution should be a risk-substitute for monotonic z.

However, because the principal cannot observe p, the state of Nature, or the agent's input use, the principal's problem is to design a contract that awards the agent for what is observable, ex post output, while ensuring the agent the reservation utility and still coming as close as possible to maximal expected social surplus. Let $\mathcal{Y}$ be the class of all functions $y : \Re_+ \to \Re$ that the principal can choose from in designing an agent reward scheme. The reward scheme works as follows: If the agent realizes an output of z then his income is set at $y(z)$ by the principal. In picking such a reward scheme, the principal must realize that if she wants to implement a particular state-contingent output vector $(z_1; z_2)$, that state-contingent output vector must be both technically feasible and consistent with the agent's private optimization in the sense that

$$\begin{aligned}(z_1; z_2) &\in \arg\max_{p;z} \{\pi_1 u(y(z_1)) + \pi_2 u(y(z_2)) - \mathrm{Min}_p \{c(z;p)\}\} \\ &= \arg\max_{z} \{\pi_1 u(y(z_1)) + \pi_2 u(y(z_2)) - C(z)\}.\end{aligned}$$

Notice, in particular, that here we use the agent's private-cost function because pollution is not observable or contractible.

The principal's problem, therefore, is

$$\max_{y \in \mathcal{Y}} \left\{ \begin{array}{c} \pi_1 (z_1 - y(z_1)) + \pi_2 (z_2 - y(z_2)) - m(p(z)) : \\ (z_1; z_2) \in \arg\max \{\pi_1 u(y(z_1)) + \pi_2 u(y(z_2)) - C(z_1; z_2)\}; \\ \pi_1 u(y(z_1)) + \pi_2 u(y(z_2)) - C(z_1; z_2) \geq \bar{u} \end{array} \right\};$$

for technically feasible z.

Under relatively weak conditions (Quiggin and Chambers, 1998), this version of the principal's problem is equivalent to designing a state-contingent payment structure subject to a set of constraints which make it privately rational for the agent to pick the desired state-contingent production structure. As discussed by Quiggin and Chambers (1998), the payment scheme operates in the following way: When the agent realizes an ex post output of, say, z she receives an ex post payment of y_1 if $z = z_1$, a payment of $z = z_2$, and an arbitrarily large negative payment otherwise. We refer to $\{(y_1, y_2), (z_1, z_2)\}$ as the contract in what follows. Formally, the principal's problem becomes:

$$\max_{(y,z)} \left\{ \begin{array}{c} \pi_1 (z_1 - y_1) + \pi_2 (z_2 - y_2) - m(p(z)) : \\ \pi_1 u(y_1) + \pi_2 u(y_2) - C(z_1, z_2) \geq \bar{u}, \\ \pi_1 u(y_1) + \pi_2 u(y_2) - C(z_1, z_2) \geq u(y_1) - C(z_1, z_1), \\ \pi_1 u(y_1) + \pi_2 u(y_2) - C(z_1, z_2) \geq u(y_2) - C(z_2, z_2) \\ \pi_1 u(y_1) + \pi_2 u(y_2) - C(z_1, z_2) \geq \pi_1 u(y_2) + \pi_2 u(y_1) - C(z_2, z_1) \end{array} \right\}.$$

The last three constraints in this problem are the ones that ensure that the agent finds it privately rational to pick the state-contingent production structure in return for the state-contingent reward structure offered by the principal.

Manipulating the second and third constraints demonstrates that contracts must be monotonic in the sense that whichever state has the highest output associated with it must also have the highest payment (Quiggin and Chambers, 1998). We state this as a lemma for future reference.

Lemma 1: Any solution must satisfy

$$(y_1 - y_2)(z_1 - z_2) \geq 0,$$

with equality only when both terms on the left-hand side are zero.

Lemma 1 implies that the principal chooses a contract with a fixed payment if and only if the contract also has a fixed production level. An example of a mechanism which would implement such a contract is a bonus-type scheme in which the agent receives no payment for production below a certain threshold and a constant payment once that threshold is reached, that is,

$$y(z) = \begin{cases} 0 & \text{if} \quad z < \hat{z} \\ \hat{y} & \text{if} \quad z \geq \hat{z}. \end{cases} \tag{2}$$

Provided that $\hat{y}$ allows the agent to reach his reservation utility level, he will accept this contract and devote all his effort to ensuring that he realizes $\hat{z}$.

Following Grossman and Hart (1983), Weymark (1986), and Quiggin and Chambers (1998), the principal's problem can be rewritten:

$$\max_{z} \left\{\pi_1 z_1 + \pi_2 z_2 - m(p(z)) - Y(z_1; z_2; \bar{u})\right\}.$$

$Y(z_1; z_2; \bar{u})$ represents what Quiggin and Chambers (1998) term the agency-cost function and is defined as the least costly way in an expected value sense for the principal to get the agent to adopt $(z_1; z_2)$ in an incentive compatible manner. $Y(z_1; z_2; \bar{u})$, thus, represents the agent's expected payment for producing $(z_1; z_2)$. Mathematically,

$$Y(z_1; z_2; \bar{u}) = \min_{y} \left\{ \begin{array}{c} \pi_1 y_1 + \pi_2 y_2 : \\ \pi_1 u(y_1) + \pi_2 u(y_2) - C(z_1; z_2) \geq \bar{u}; \\ \pi_1 u(y_1) + \pi_2 u(y_2) - C(z_1; z_2) \geq u(y_1) - C(z_1; z_1) \\ \pi_1 u(y_1) + \pi_2 u(y_2) - C(z_1; z_2) \geq u(y_2) - C(z_2; z_2) \\ \pi_1 u(y_1) + \pi_2 u(y_2) - C(z_1; z_2) \geq \pi_1 u(y_2) + \pi_2 u(y_1) - C(z_2; z_1) \end{array} \right\}.$$

The solution to the agency-cost problem is one of the main contributions of our earlier paper (Quiggin and Chambers 1998). While our derivation there is relatively complicated, the end result is that the underlying moral hazard problem can be reduced to a remarkably simple nonlinear program. We do not go into details here, but instead refer the reader to our earlier work.[12] Applying Results 1- 6 of Quiggin and Chambers (1998), the principal's problem can be written

$$\max_{z} \left\{\pi_1 z_1 + \pi_2 z_2 - m(p(z)) - Y(z_1; z_2; \bar{u}) : z_2 \geq z_1\right\};$$

where

$$(3) \qquad Y(z_1; z_2; \bar{u}) = \pi_1 h(\bar{u} + C(z_1; z_1)) + \pi_2 h\left(\bar{u} + C(z_1; z_1) + \frac{C(z_1; z_2) - C(z_1; z_1)}{\pi_2}\right);$$

and h is the inverse mapping of u. A brief comment about this result is appropriate before proceeding. Under (1) and homogeneity of C, the optimal state-contingent production structure must be monotonic.[13] In other words, the principal will order states of Nature exactly as would a risk-neutral individual facing the technology. Because the production structure is

[12] Chambers and Quiggin (2000, Chapter 9) also gives a complete treatment of the problem.

[13] This follows from Result 1 and Lemma 4 in Quiggin and Chambers (1998).

monotonic in this sense, Lemma 1 then implies that the state-contingent payment structure must be as well. Second, for given z, the optimal state-contingent payments are given by

$$y_1 = h\left(\bar{u} + C(z_1; z_1)\right) \tag{4}$$

and

$$y_2 = h\left(\bar{u} + C(z_1; z_1) + \frac{C(z_1; z_2) - C(z_1; z_1)}{\pi_2}\right). \tag{5}$$

Intuitively, for higher-powered state-contingent contracts, the reward schedule relating payments to output (which by Lemma 1 is positively sloped) is in some sense steeper than for lower-powered incentives. For example, in comparing two affine payment schedules with a common intercept, the one with the steeper slope has the higher powered incentives. Now consider comparing two affine payment schedules with a common slope but a different intercept. The schedule with the higher intercept now involves higher overall incentives for all ex post outputs. The certainty of higher income associated with it will entice some agents to accept it who would reject the other contract. If we compare two affine payments schemes with different intercepts and slopes, the intercepts can be chosen so that the one with the steeper slope offers lower returns for all relevant ex post outputs than the alternative. More generally, for nonlinear contract structures, the difference between a higher-powered and a lower-powered contract involves more than just the slope of the payment schedule for risk-averse agents.[14]

Alternatively, one can also observe that the contract with what are traditionally viewed as the highest-powered incentives, where the agent is made the residual claimant, exposes the agent to the entire spectrum of income and production risk. Moreover the contract with what are usually perceived as the lowest-powered incentives, where the agent receives a fixed payment regardless of the production outcome, exposes the agent to no income risk. Therefore, we base our formal definition of higher powered incentives upon the relative dispersion of contracts. Contract $\{y^A; z^A\}$ has higher powered incentives than $\{y^B; z^B\}$ if

$$\begin{aligned} y_2^A &\geq y_2^B \geq y_1^B \geq y_1^A \\ z_2^A &\geq z_2^B \geq z_1^B \geq z_1^A. \end{aligned}$$

Thus, our notion of higher-powered contracts is that they are 'spread out' versions of lower-powered contracts. The agent is asked to produce a more dispersed output distribution in return for a more dispersed income payment distribution.

[14] Both Holmström and Milgrom (1991) and Baker (1992) use linear payment schedules. In Holmström and Milgrom (1991), this ambiguity is removed by restricting attention to individuals with constant absolute risk aversion. In Baker (1992), individuals are risk neutral.

3 Optimality of Fixed Payment Schemes

By Lemma 1, fixed payments only emerge in concert with fixed production. Our notion of a fixed payment-production standard is the lowest-powered incentive scheme possible, where the optimal contract can be implemented by a mechanism of the form (2). Both the state-contingent production and the payment are fixed. For mnemonic simplicity, we refer to the case where the contract does not involve a fixed payment as an incentive contract. Suppose that $z_1 = z = z_2$, is optimal. By (3),

$$Y(z;z;\bar{u}) = h(\bar{u} + C(z;z));$$

with $h(\bar{u} + C(z;z))$ representing the fixed payment to the agent. The optimal value of the principal's objective function is then

$$z - h(\bar{u} + C(z;z)) - m(p(z;z)):$$

The only implementable alternatives to the fixed payment-production standard are strictly monotonic contracts with $z_2 > z_1$. Consider the multiplicative spread of the output vector defined by increasing z_2 by the small positive amount δz_2 and decreasing z_1 by $-\frac{\pi_2}{\pi_1}\delta z_2$. The induced change in the principal's welfare has the same sign as

$$\text{(6)} \qquad -h'(\bar{u} + C(z;z))\left(\frac{C_2(z;z)}{\pi_2} - \frac{C_1(z;z)}{\pi_1}\right) - m'(p(z;z))\left(\frac{p_2(z;z)}{\pi_2} - \frac{p_1(z;z)}{\pi_1}\right);$$

where $'$ denotes a derivative. If $p(z;z)$ is a risk substitute, moving to a riskier z decreases p, whence $\frac{p_2(z;z)}{\pi_2} - \frac{p_1(z;z)}{\pi_1} < 0$. Thus, by (1) such an increase in the riskiness of z unambiguously increases the principal's welfare in this case. Hence, a fixed payment-production standard cannot be optimal if $p(z;z)$ is a risk substitute, and the multiplicative spread is implementable. For a fixed payment-production standard to be optimal, such an increase in the riskiness of z must make the principal's welfare fall. By (1), this can only happen if $\frac{p_2(z;z)}{\pi_2} - \frac{p_1(z;z)}{\pi_1} \geq 0$. But this requires that $p(z)$ rises as a result of the move to the riskier z.

Proposition 1: If $p(z)$ is a risk substitute, and the principal can implement an incentive scheme, a fixed payment-production standard cannot be optimal.

A word about the structure of our problem is appropriate. The unmeasurable task, p, has been modelled as a 'bad'. Results change in a predictable manner if the unmeasurable task is a 'good'. In that case, the multiplicative spread above would increase the principal's expected profit from production as expected return does not change, but agent cost falls by

(1). Hence, if the unmeasurable task is a good, then a fixed payment is optimal only if it is a risk substitute. A similar argument applies to the remaining results in the paper, and we leave it to the reader to appropriately adjust our stated results.

Proposition 1 may be given some heuristic content by considering the case where the agent is a factory employee, and p is wear and tear on a piece of capital equipment that is used by several employees. It seems plausible that p will be directly related to the rate and intensity at which the equipment is operated. Suppose that the equipment is such that if all things go well, its productivity in terms of z increases dramatically when it is operated at high speed but diminishes even more rapidly if things go poorly. Here one can think of the states of Nature as being the worker's surrounding environment, his health, etc. Then in our terms, p would be risk complementary. And Proposition 1 suggests that the firm might consider paying workers a straight wage rather than a piece rate. Piece rates, in this instance, could lead to excessive wear and tear on the capital stock.

Proposition 1 remains true even if (1) holds as an equality. In that instance, the multiplicative spread of z leaves the agent's cost, and hence the principal's agency cost, unchanged. But if $p(z)$ is a risk substitute, then the riskier z still carries with it a lower p; which is welfare improving. Chambers and Quiggin (1997,2000) have defined a class of technologies for which (1) always holds as an equality– the generalized Schur convex technologies. C is generalized Schur convex if

$$z \succeq_{\pi} z' \Rightarrow C(z) \leq C(z'); \quad z \in \Re^2_+.$$

Intuitively, it is the class of stochastic technologies which leads a risk-neutral individual to always pick a non-stochastic production vector.[15]

Quiggin and Chambers (1998, Corollary 8.1) show that a principal facing single-task moral hazard with an agent using a generalized Schur convex technology can completely eliminate that moral hazard. If costs are generalized Schur convex, there is no tension between risk allocation and production efficiency. Production efficiency then requires producing a non-stochastic output which, in turn, is implementable by a fixed payment.

In the multi-task case, the principal's ability to resolve the agency problem must be traded off against her ability to control $p(z)$ indirectly through the contract. Thus, even when the technology is generalized Schur convex, the principal prefers an incentive contract if p is a risk substitute.

[15] As Chambers and Quiggin (1997) demonstrate, the non-stochastic technology is a degenerate special case of the class of generalized Schur convex technologies. However, generalized Schur convex technologies are not trivially stochastic as also demonstrated by Chambers and Quiggin (1997).

Corollary 1.1: If C is generalized Schur convex, and if $p(z)$ is a risk substitute at z optimal for the single-task moral-hazard problem, the principal prefers a multiplicative spread of z to z.

Corollary 1.1 and Lemma 1 imply that the principal always gains, as compared to the single-task optimal contract, by introducing a riskier production structure. Hence, when costs are generalized Schur convex, the optimal regulatory scheme will involve higher-powered incentives for production than in the single-task case if $p(z)$ is a risk substitute.

Corollary 1.1 establishes a seemingly paradoxical result. Even when it is privately cheapest to implement a non-stochastic z and doing so completely removes agency costs, the principal's concerns about p will lead him to insist on the agent deploying a stochastic production vector. Multi-tasking causes the principal to insist on an incentive scheme where he would otherwise implement a standard.

If there is to be a case made for standards when incentive schemes are feasible, it must rest on the presumption that the noncontractible activities are risk-complementary. In fact, it is an easy consequence of preceding arguments to state a sufficient condition for standards to be optimal.

Proposition 2: A fixed payment-production standard is always optimal if C is generalized Schur convex and $p(z)$ is risk-complementary.

When the cost structure is generalized Schur convex, the incentive problems associated with the presence of moral hazard can be efficiently surmounted by implementing a production standard. If $p(z)$ is risk-complementary as well, then the principal's evaluation of the damage caused by the unobservable task also pushes him toward a standard.

There are, however, instances where incentive problems are so severe that even in the presence of risk-substitutability, they will prevent the principal from implementing anything other than a production standard. One such limiting case is given by the case of complete aversion to risk[16] where the agent's attitudes toward state-contingent incomes are given by

$$W(y_1, y_2) = \min\{y_1, y_2\}. \tag{7}$$

In an appendix, we show

Lemma 2 If agent preferences are given by (7), z is implementable if and only if $z_1 = z_2$.

[16] Of course, this case departs from the expected-utility hypothesis that underlies the rest of the paper. However, we can arrive at this characterization by assuming constant absolute risk aversion and allowing the risk parameter to go to infinity.

Because the agent is so averse to risk, she only cares about the lowest payment she may receive. Thus, offering her a reward for realizing a higher output in one of the states of Nature does not elicit greater effort. Ex ante, it carries no marginal gain in welfare. All her effort is concentrated on ensuring that she realizes the lowest acceptable output. The principal's problem, therefore, becomes one of picking the optimal standard. A natural conjecture, therefore, is that as the agent's risk aversion increases, the principal becomes less likely to rely upon a high powered contract because the cost of implementing such contracts grows with the agent's risk aversion.

By Lemma 2, the agency cost function is now $C(z;z) + \mu$. The principal's problem is to

$$\max_z \{z - C(z;z) - m(p(z;z))\} - \mu:$$

The solution is straightforward, and therefore we do not discuss it in detail. Even here, however, a multiplicative spread about the optimal z improves the principal's welfare if $p(z)$ is a risk substitute. But as Lemma 2 establishes, this riskier z is never implementable.

Corollary 1.2: If agent preferences are given by (7), ignoring incentive effects, the principal always prefers an incentive scheme to a standard if $p(z)$ is a risk substitute.

Optimality of fixed payments and production standards requires the presence of risk-complementarity. One can go further. Defining an auxiliary variable, α, by

$$z_2 = z_1 + \alpha;$$

the principal's first-order conditions for the multi-task problem are

$$\begin{aligned} 1 - Y_1(z_1;z_2;\mu) - Y_2(z_1;z_2;\mu) - m'(p(z))(p_1(z) + p_2(z)) &\leq 0; \quad z_1 \geq 0; \\ \pi_2 - Y_2(z_1;z_2;\mu) - m'(p(z))\,p_2(z) &\leq 0; \quad \alpha \geq 0; \end{aligned} \tag{8}$$

in the notation of complementary slackness. We restrict attention to the case where a production standard is optimal at a positive level of production. Hence, the first expression in (8) requires

$$Y_2(z;z;\mu) + m'p_2(z;z) = 1 - Y_1(z;z;\mu) - m'p_1(z;z): \tag{9}$$

A standard is only optimal if the optimal solution requires $\alpha = 0$, whence the second expression in (8) requires

$$\pi_2 - Y_2(z;z;\mu) - m'p_2(z) < 0: \tag{10}$$

Adding (9) and (10) yields

$$\pi_1 - Y_1(z;z;\theta) - m'p_1(z;z) > 0. \tag{11}$$

Recognizing that in the case of a standard $Y(z;z;\theta) = h(\theta + C(z;z))$, allows us to rewrite (10) and (11) as

$$\begin{aligned} \pi_1 - h'(\theta + C(z;z))\,C_1(z;z) - m'p_1(z;z) &> 0, \\ \pi_2 - h'(\theta + C(z;z))\,C_2(z;z) - m'p_2(z;z) &< 0. \end{aligned} \tag{12}$$

The left-hand side of the inequalities in expressions (12) can be recognized as the derivatives of the principal's objective function in the case where he is able to observe the state of Nature, but p remains unobservable. Because the state of Nature is observable, the principal can write enforceable state-contingent contracts without fear of providing adverse incentives. Hence, it is optimal to offer the agent a fixed payment equalling $h(\theta + C(z))$. However, because the principal cannot observe p, he cannot regulate it by charging an appropriate Pigouvian tax or granting a subsidy. The agent's private cost function and $p(z)$, therefore remain relevant. We shall refer to this as the second-best problem and note that it corresponds with the classic externality problem.[17] For the sake of simplicity, assume that the principal's second-best objective function is concave in z.

Expressions (12) require that the solution to the second-best problem involves a non-monotonic production structure. Thus, a fixed payment-production standard is optimal only if p is a strong enough risk complement to force the principal to overturn the sorting of states of Nature imposed by (1). With moral hazard, a non-monotonic z is not implementable. The closest the principal can come to the second-best non-monotonic production structure is to offer the agent a fixed payment and a production standard.

For the converse, suppose that the principal's solution to the second-best problem is non-monotonic. Call it z^{SB}. For any monotonic z, there is a $\lambda \in [0,1]$ such that $z^* = \lambda z^{SB} + (1-\lambda)z$ lies on the bisector. By the definition of z^{SB}

$$E_\pi[z^{SB}] - h(C(z^{SB}) + \theta) - m\left(p(z^{SB})\right) \geq E_\pi[z] - h(C(z) + \theta) - m(p(z)), \tag{13}$$

where E_π denotes the expectations operator with respect to π. Therefore:

$$\begin{aligned} E_\pi[z^*] - Y(z^*;\theta) - m(p(z^*)) &= E_\pi[z^*] - h(C(z^*) + \theta) - m(p(z^*)) \\ &\geq E_\pi[z] - h(C(z) + \theta) - m(p(z)) \\ &\geq E_\pi[z] - Y(z;\theta) - m(p(z)) \end{aligned}$$

[17] In earlier versions of this paper, we referred to the second-best problem as the externality with insurance problem.

The equality follows from (3) because z^* lies on the bisector. The first inequality follows by concavity of the principal's objective function and (13). The final inequality follows from Result 2 of Quiggin and Chambers (1998), which establishes that $Y(z;\bar{u})$ is always bounded below by $h(C(z)+\bar{u})$.[18] Combining results, we have established that in this case there always exists a production standard which dominates any implementable monotonic contract. Thus, we conclude:

Proposition 3: If $C(z_1; z_2)$ is strictly monotonic, a fixed payment-production standard is optimal if and only if the solution to the second-best problem has non-monotonic z.

Our notion of a fixed payment-production standard involves the principal specifying a contract where the agent produces a nonstochastic output in return for a nonstochastic payment. Holmström and Milgrom (1991), however, have identified conditions under which a principal finds it optimal to offer the agent a non-stochastic reward in return for a stochastic output. By Lemma 1, this cannot happen in our model. The reason lies in our assumption of strict monotonicity of the agent's private cost structure. If that assumption is relaxed, then incentive compatibility only requires that the payment structure be weakly monotonic. The possibility of a fixed payment for a stochastic output then emerges. However, as we show immediatly below, fixed payments typically will be optimal for cost structures that are not strictly monotonic, regardless of the presence of multi-task concerns, unless production of z is technically inefficient. Original specifications of agency problems in terms of a stochastic production function (e.g., Ross, 1973; Harris and Raviv, 1979) presumed that production was always technically efficient. Therefore, in addition to identifying how our results change as a consequence of relaxing the strict monotonicity assumption, the discussion in the next section also helps explain some of the difficulty early moral-hazard studies encountered in identifying monotonic contracts.

3.1 Weak monotonicity

Thus far, we have assumed that the cost structure is strictly monotonic. Chambers and Quiggin (2000, Chapter 4) have shown that specifying the state-contingent technology by a stochastic production function yields a cost structure that is only weakly monotonic. More precisely, cost is non-decreasing in z; but is strictly increasing only along a one-dimensional

[18] Because Quiggin and Chambers (1998) only consider a single-task moral hazard problem, $h(C(z)+\bar{u})$ is what they refer to as the first-best agency cost function, Y_{FB}. This referenced result, therefore, simply reflects the fact that contracts with incentives are more expensive for the principal to implement than the first-best contract. Note that this observation underlies much of the empirical work on incentive contracts. Prendergast (1999) contains a review of this literature.

expansion path determined by the stochastic production function. Quiggin and Chambers (1998) and Chambers and Quiggin (2000, Chapter 9) pinpoint this weak monotonicity as the stumbling block to establishing contract monotonicity in the original state-space formulations of the principal-agent problem. The following cost structure

$$C(z_1, z_2) = \max\left\{c^1(z_1), c^2(z_2)\right\},$$

corresponds to a stochastic production function satisfying free disposability of state-contingent output (Chambers and Quiggin 2000, Chapter 4), and will be referred to as weakly monotonic. Here $c^i(z_i)$ is the minimal ex post cost associated with producing z_i. Imposing positive linear homogeneity on this specification requires

$$c^i(z_i) = c_i z_i$$

with $c_i > 0$. To ensure consistency with (1) in this case where $C(z_1, z_2)$ is not smoothly differentiable, assume that the first state of Nature is the 'bad' state of nature so that $c_1 > c_2$. Regardless of their risk attitudes or their subjective probabilities, residual claimants facing this technology always choose $z_2 > z_1$.

An optimal solution to the single-task problem for this cost structure is for the principal to offer a fixed payment, subject to the production of some minimum level of output z_1. For any monotonic $z = (z_1, z_2)$ with $c_1 z_1 \geq c_2 z_2$ consider the fixed payment

$$y(z_1) = y(z_2) = h(\bar{u} + c_1 z_1).$$

Because $C(z_1, z_1) = C(z_1, z_2) = c_1 z_1$, this contract offers the agent exactly his reservation utility. It is also incentive compatible because for monotonic z, $C(z_2, z_2) = c_1 z_2 \geq c_1 z_1$. If $c_1 z_1 < c_2 z_2$, the principal can achieve a higher level of z_1 with no change in cost to the agent.

Output vectors z with $c_1 z_1 < c_2 z_2$ can be implemented with a payment vector (y_1, y_2) satisfying

$$\begin{aligned} \pi_1 u(y_1) + \pi_2 u(y_2) - c_2 z_2 &= \bar{u} \\ \pi_1 u(y_1) + \pi_2 u(y_2) - c_2 z_2 &= u(y_1) - c_1 z_1. \end{aligned}$$

However, for any such z, there exists z' with

$$\begin{aligned} C(z_1', z_2') &= c_1 z_1' = c_2 z_2' \\ z_1 &< z_1' < z_2' < z_2 \\ \pi_1 z_1' + \pi_2 z_2' &= \pi_1 z_1 + \pi_2 z_2, \end{aligned}$$

which can be implemented with payment structure

$$y(z_1') = y(z_2') = h(\bar{u} + C(z_1'; z_2')).$$

Since

$$C(z_1; z_2) = c_2 z_2 > c_2 z_2' = C(z_1'; z_2'),$$

we have, by the convexity of h that

$$\begin{aligned} \pi_1 y_1 + \pi_2 y_2 &> h(\bar{u} + c_2 z_2) \\ &> h(\bar{u} + C(z_1'; z_2')). \end{aligned}$$

The contract associated with z' thus offers the principal the same expected return but a lower cost. Hence, he will always prefer the fixed-payment contract eliciting output z' to the incentive-based contract eliciting z.

Now consider the general multitask case. As in the single-task case, (4) and (5) imply that $y_2 \geq y_1$ if and only if

$$\max\{c_1 z_1; c_2 z_2\} = C(z_1; z_2) \geq C(z_1; z_1) = c_1 z_1.$$

The optimal payment structure will thus be strictly monotonic if and only if

$$c_2 z_2 > c_1 z_1. \tag{14}$$

Expression (14), however, requires that the state-contingent output bundle be technically inefficient in the sense that the agent could produce strictly more z_1 while maintaining the level of z_2 without incurring any higher cost. Hence, for this cost structure, we conclude on the basis of these arguments that

Proposition 4: For the weakly monotonic cost structure

$$C(z_1; z_2) = \max\{c_1 z_1; c_2 z_2\},$$

the principal offers strictly monotonic payments if and only if production of z is technically inefficient. In the single-task problem, the principal always offers a fixed payment, and production of z is never technically inefficient.

Clearly, the only situation where the principal will desire a technically inefficient production pattern is one in which $p(z)$ is strongly risk-substituting, so that the benefits which arise in the single-task problem from the adoption of the less risky output z' in place of z are offset by an increase in $p(z)$.

4 Optimal Contracts and the Structure of Incentives

We now turn to the broader issue of whether multi-task problems lead to higher powered or lower powered incentives. We use the single-task moral hazard problem as a point of reference. The single-task optimum yields a particularly interesting benchmark because it corresponds to the case where the principal effectively delegates the responsibility for the choice of p completely to the agent.

4.1 The Single-Task Optimum

We only consider the case where a standard is not optimal in the single-task case.[19] The first-order conditions for an optimum then are

$$
\begin{aligned}
\pi_1 - Y_1(z_1; z_2; \theta) &\leq 0; \quad z_1 \geq 0; \\
\pi_2 - Y_2(z_1; z_2; \theta) &= 0; \quad z_2 > 0;
\end{aligned}
$$

in the notation of complementary slackness. Quiggin and Chambers (1998) fully characterize the solution to this problem. The interested reader can refer to that paper for details. For an interior solution,

$$
\frac{Y_2(z_1; z_2; \theta)}{\pi_2} = \frac{Y_1(z_1; z_2; \theta)}{\pi_1}: \tag{15}
$$

Observing this solution for the single-task moral hazard problem allows us to state the following lemma (proof is in an appendix):

Lemma 3 At the optimal solution to the single-task moral hazard problem, implementing a (small) multiplicative spread of z yields a higher powered contract.

In the single-task moral hazard problem, a move to a higher powered contract has no impact on the principal's welfare. Hence, the principal obviously has no incentive to move in this direction, but as the next section illustrates, matters change in the multi-task case.

4.2 Higher vs. Lower Powered Contracts

Evaluate the principal's objective function at the solution to the single-task agency problem and consider a mean-preserving increase in the riskiness of z. The resulting change in the principal's objective function is

$$
\left[-\left(Y_2(z_1; z_2; \theta) - \frac{\pi_2}{\pi_1} Y_1(z_1; z_2; \theta) + m'(p(z)) \left(p_2(z) - \frac{\pi_2}{\pi_1} p_1(z) \right) \right) \right] \delta z_2
$$

[19] This implies that $z_2 > 0$:

with $\delta z_2 > 0$. By (15), this reduces to

$$-m'(p(z))\left(p_2(z) - \frac{\pi_2}{\pi_1}p_1(z)\right)\delta z_2.$$

Lemma 3, in turn, implies that such a mean-preserving increase in the riskiness of z leads to a mean-preserving increase in the riskiness of y. On this basis, we conclude

Proposition 5: At the single-task agency optimum, the principal benefits from a contract structure with higher-powered incentives if $p(z)$ is a risk substitute. At the single-task agency optimum, the principal benefits from a contract structure with lower-powered incentives if $p(z)$ is risk-complementary.

Proposition 5 confirms our intuition. Suppose that the principal ignores the possibility of multi-tasking. He then offers the agent a contract that trades off efficiency losses against enhanced provision of incentives. It is often alleged that contracts which focus solely on measurable outcomes provide too high powered incentives and encourage the agent to divert attention from other tasks toward producing the contracted outcomes. Baker (1992), for example, contains a clear illustration of this effect for risk-neutral agents. Proposition 5 qualifies this argument by showing that the principal can benefit from providing even higher powered incentives when $p(z)$ is a risk substitute. However, when $p(z)$ is a risk complement, the appropriate response is for the principal to mute the incentives associated with the contracted outcomes, as in the case examined by Baker.

As an illustration of this effect, consider the case of preventive medicine. Suppose that p is the cost of tests and diagnostic procedures, and that doctors' contractual arrangements cannot be made contingent on p. Then if doctors face high-powered incentives, for example in the form of potentially bankrupting malpractice suits, and if diagnostic procedures reduce the riskiness of patient outcomes, excessively high values of p will be chosen. Insurers or governments providing health funding will therefore have an incentive either to ensure that doctors face lower-powered incentives (for example, through tort law reform) or to design contracts in which p can be monitored (for example, managed care contracts).

5 An application to Nonpoint-Source Pollution Control

Earlier, we briefly discussed a specific multi-task, principal-agent problem: the design of optimal crop payments schemes in the presence of noncontractible runoff pollution. Government supported payment schemes, typically advertised as protection against market vagaries, have a long standing in both the United States and the European Community. In recent years, it has become increasingly recognized that such schemes should be designed taking into account the potential adverse effects that agricultural production has on the environment (Lichtenberg and Zilberman, 1986). Thus, in recent U.S. farm bills there has been increased use of environmental provisions such as 'swamp buster' clauses, conservation acreage reserves, and the installation of riparian buffers. Such programs are typically based on observable conservation practices, such as the planting of fragile lands to cover crops instead of commercial crops. In a broad sense, therefore, the optimal design of these policies can be addressed by standard Pigouvian analysis.

Our model is more apposite to the case where the environmental activity is not contractible but is directly related to the production activity. Such is the case with the regulation of non-point source pollution from farming in the form of runoff. Segerson (1988) studied the optimal regulation of nonpoint-source pollutants for an agent who is indifferent to the dispersion of runoff. Dosi and Tomasi (1994) contains a number studies on the optimal regulation of nonpoint-source pollution control under different informational and institutional settings including hidden action and hidden information. None of these studies, however, address the problem in a multi-task framework with moral hazard. Chambers and Quiggin (1996) have studied nonpoint-source pollution control as a multi-task problem under different informational assumptions than used here. Specifically, they consider pure hidden action in the sense that the agent's effort, crop output, and runoff level are all unobservable. What is observable and contractible is the state of Nature. Therefore, unlike the present model, hidden action is not coupled with hidden knowledge.

In this example, - corresponds to climatic conditions that are beyond the farmer's control. z represents crop output in the two states of Nature, and p is runoff pollution from farming, which is assumed unobservable by the principal who is now taken to be a social planner. Thus, under our assumptions one might think intuitively of state 1 as indexing relatively low natural moisture or drought conditions, while state 2 corresponds to adequate natural moisture. Or alternatively, state 1 might index a severe natural pest infestation and state 2 might index no pest infestation. We are particularly interested in studying the possibility of a command and control approach to the regulation of nonpoint-source runoff pollution. In actual policy experience, these are often manifested in the form of acreage or production quotas or mandated 'best management practices'. If p is a risk substitute, these practices are likely not optimal. Therefore, we initially focus attention on the case where runoff is risk-complementary.

Assume that the agent's cost function can be written as

$$c(z;p) = c_1 z_1 + c^2(z_2;p);$$

where $c_1 > 0$ and $c^2(z_2;p)$ is positively linearly homogeneous in its arguments and strictly increasing in z_2. This particular technology is a special case of what Chambers and Quiggin (1996, 2000) refer to as a state-allocable technology. Specifically, it implies that the agent's tasks can be targetted at two production activities: preparing the first state-contingent output and preparing the second state-contingent output and the runoff. Runoff, therefore, is most naturally associated with the production of the second state-contingent output. This assumption is made solely to streamline the analysis and to allow for the statement of clear-cut results.

In general, one expects runoff to be associated with both state-contingent outputs. The case of chemical fertilizer illustrates. If applied with too little moisture from natural or man-made sources, chemical fertilizer can sharply reduce yield. However, if moisture is adequate, its application can greatly enhance yield. Thus, assuming runoff is directly associated with chemical application, one expects there to be regions in which chemical fertilizer is risk complementary and is related to both z_1 and z_2. Our specification, therefore, would perhaps be most appropriate for the case where 1 indexes natural moisture that is too little for optimal growth but sufficient to avoid chemical 'burn' of the crop.[20]

The farmer's private cost function is now

$$\begin{aligned} C(z) &= c_1 z_1 + \mathrm{Min}_p \left\{ c^2(z_2;p) \right\} \\ &= c_1 z_1 + z_2 \mathrm{Min}_{\frac{p}{z_2}} \left\{ c^2\left(1; \frac{p}{z_2}\right) \right\} \\ &= c_1 z_1 + z_2 c^2(1;\rho) \\ &\equiv c_1 z_1 + c_2 z_2. \end{aligned} \tag{16}$$

where

$$\rho = \arg\min_{\frac{p}{z_2}} \left\{ c^2\left(1; \frac{p}{z_2}\right) \right\}:$$

Unobservable runoff is, therefore, given by ρz_2. Because runoff is always proportional to z_2 and unrelated to z_1, it is risk-complementary in a very strong sense. Regardless of what

[20] In the agricultural-economics literature, the concept of a risk increasing input corresponds approximately to our notion of risk complementary. Chemical fertilizers are cited as an example of a risk increasing input. Recently, there has been an empirical debate on whether fertilizers, and in particular chemical fertilizers, are in fact risk increasing (Horowitz and Lichtenberg, 1993; Babcock and Hennessy, 1996).

happens to the dispersion of z, runoff increases if z_2 increases. This specification, therefore, allows for the possible optimality of production standards as a means of controlling the runoff.

The agent's private isocost curve, therefore, corresponds to a negatively sloped line segment in state-contingent output space. Under (1), a risk-neutral individual facing this technology would, therefore, set $z_1 = 0$.

Assume that the social marginal cost of runoff is a positive constant μ. The planner's second-best problem in this framework, therefore, can be written

$$\mathrm{Max}_z \{\pi_1 z_1 + \pi_2 z_2 - \mu\rho z_2 - h(\bar{u} + c_1 z_1 + c_2 z_2)\};$$

and the associated first-order conditions, which are necessary and sufficient, are

$$\begin{aligned} \pi_1 - h'(\bar{u} + c_1 z_1 + c_2 z_2)\, c_1 &\le 0; \quad z_1 \ge 0 \\ \pi_2 - \mu\rho - h'(\bar{u} + c_1 z_1 + c_2 z_2)\, c_2 &\le 0; \quad z_2 \ge 0; \end{aligned}$$

in complementary slackness notation.

Any solution to the second-best problem must, therefore, satisfy

$$h'(\bar{u} + c_1 z_1 + c_2 z_2) \ge \max\left\{\frac{\pi_1}{c_1}, \frac{\pi_2 - \mu\rho}{c_2}\right\}. \tag{17}$$

$\frac{\pi_1}{c_1}$ and $\frac{\pi_2 - \mu\rho}{c_2}$, respectively, are the social benefit cost ratios associated with a marginal increase in z_1 and z_2, while $h'(\bar{u} + c_1 z_1 + c_2 z_2)$ represents the marginal cost of increasing the agent's welfare. Because of the linear nature of cost and damages, optimality requires setting the latter equal to the maximal of the two former quantities. The optimal solution involves setting $z_1 = 0$, $z_2 = 0$, or is indeterminate. We assume away the indeterminate solutions in what follows..

Therefore, it is a straightforward consequence of Proposition 3 that:

Corollary 3.1: In the nonpoint-source pollution control model, a fixed payment-production standard is optimal only if $\frac{\pi_1}{c_1} > \frac{\pi_2 - \mu\rho}{c_2}$.

A production standard, therefore, is only optimal if the damage associated with runoff from crop production is so large that it would lead even a risk-neutral producer (forced to internalize the cost of runoff) to reverse the ordering of states dictated by (1). In fact, environmental damage is so large that it is second-best optimal for the producer only to produce a positive amount in the 'bad' state of nature. No effort would then be devoted to z_2, and consequently run off would be zero as well. However, this outcome is not monotonic and, therefore, is not implementable. The planner thus opts for the next best thing which is a standard.

There are a number of obvious and intuitive corollaries to Corollary 3.1. They include that production standards are more likely to be an optimal way of regulating the nonpoint-source pollutant the larger is the rate at which runoff occurs and the larger is the marginal damage associated with runoff. Standards are less likely to be optimal, even in the case of strong risk-complementarity, the larger is the initial gap between the marginal-cost probability ratios given in (1).

For the remainder of this section, we will consider the case where runoff is a strong risk complement, but production standards are not optimal. Our goal, of course, is to ascertain the effect that the presence of multi-task considerations has on the design of an optimal crop-payment scheme. Therefore, maintain the assumption that $\frac{\pi_2 - \rho}{c_2} > \frac{\pi_1}{c_1}$. The optimal second-best solution sets $z_1 = 0$. Denote the optimal level of z_2 by z_2^{SB} and the associated optimal non-stochastic payment by $y^{SB} = h\left(\bar{u} + c_2 z_2^{SB}\right)$.

We first consider the solution to the single-task agency problem in which there is no pooling of payments. Substituting the assumed cost structure into (4) and (5), the principal's objective function can be written

$$\pi_1 z_1 + \pi_2 z_2 - \pi_1 h(\bar{u} + (c_1 + c_2) z_1) - \pi_2 h\left(\bar{u} + (c_1 + c_2) z_1 + \frac{c_2}{\pi_2}(z_2 - z_1)\right):$$

Introducing an auxiliary variable as before, the associated necessary and sufficient first-order conditions for an interior solution with no pooling of payments are

$$1 - (c_1 + c_2)\left[\pi_1 h'(u_1) + \pi_2 h'(u_2)\right] = 0; \tag{18}$$

$$\pi_2 - h'(u_2) c_2 = 0; \tag{19}$$

where $u_i = u(y_i)$. Expression (18) requires that the principal exhaust all opportunities for raising profit by having the agent nonstochastically increase production, while expression (19) implies that the principal will design the contract so that the agent chooses z_2 efficiently at the margin.

If the crop-payment contract is strictly monotonic and does not admit a production standard Lemma 1, (18), and (19) ensure that

$$\frac{\pi_2}{c_2} = h'(u_2) > \pi_1 h'(u_1) + \pi_2 h'(u_2) = \frac{1}{c_1 + c_2} \tag{20}$$

in the single-task optimum. Expression (20) says if a production standard is not optimal, the benefit cost ratio from raising z_2 must be greater than the benefit cost ratio associated with nonstochastically increasing production. If it were not, at the margin the principal is better off asking the farmer to increase production nonstochastically than asking the farmer to increase z_2. Under such circumstances, the principal would always benefit more from a production standard than from a separating solution.

Turning to the multiple-task agency problem, the principal's necessary and sufficient first-order conditions (after introducing an auxiliary variable as before) essentially repeat (18) and (19) while accounting for the presence of runoff in the design of the crop-payment schedule:

$$1 - (c_1 + c_2)\left[\pi_1 h'(u_1) + \pi_2 h'(u_2)\right] - \mu\rho \le 0; \quad z_1 \ge 0; \tag{21}$$

$$\pi_2 - h'(u_2)\, c_2 - \mu\rho = 0. \tag{22}$$

Expression (22) is written as an equality because our assumptions guarantee that a production standard is not optimal.

We first compare the solution of the multiple-task moral hazard problem to the second-best problem above. Denoting the solutions to the multiple-task problem by superscript M, we have (the proof is in the Appendix):

Proposition 6: Assume $\frac{\pi_2 - \mu\rho}{c_2} > \frac{\pi_1}{c_1}$; then for an interior solution to the multiple-task moral hazard problem,

$$y_2^M = y^{SB} > y_1^M;$$

$$C\left(z_1^{SB}; z_2^{SB}\right) > C\left(z_1^M; z_2^M\right);$$

and

$$z_2^{SB} > z_2^M > z_1^M > z_1^{SB} = 0.$$

When compared to the regulatory scheme that emerges in the second best, the multi-task regulatory scheme is characterized by four things: a lower level of effort by the agent (as measured by his private cost), a lower level of runoff, a more dispersed payment schedule, and a less dispersed production schedule.

In the second-best, it is optimal for the principal to ask the agent to produce, in effect, the riskiest type of production bundle, one that involves no output at all in the 'bad' state of nature in return for a non-stochastic payment. If the principal tries to implement such a regulatory scheme in the presence of moral hazard, the agent will always respond by committing no effort (thus ensuring no output and no runoff pollution) and simply taking the non-stochastic payment. Therefore, the principal has to simultaneously reduce the riskiness of the production bundle while exposing the agent to more risk in the payment schedule. The move away from specialization in the 'good' state output forces the principal to forego some expected output, which in turn allows the agent to reduce his cost. It also follows immediately from Proposition 6 that less runoff is emitted in the multi-task case than in the second-best case. This decrease in runoff is directly linked to the diminished effort and nonspecialization of the agent in the good state output.

It is a straightforward consequence of Proposition 4 that at the solution to the single-task moral hazard problem, the principal under the current assumptions would gain by introducing lower-powered incentives. We leave this to the reader. Our next result compares the single-task and multiple-task solutions directly. Denoting the solution to the single-task moral hazard problem by a superscript S, we have (the proof is in an appendix)

Proposition 7: Assume $\frac{\pi_2 - \mu\rho}{c_2} > \frac{\pi_1}{c_1}$; then for an interior solution to the multiple-task moral hazard problem,

$$
\begin{aligned}
y_2^S &> y_2^M > y_1^M > y_1^S; \\
z_2^S &> z_2^M > z_1^M > z_1^S.
\end{aligned}
$$

When compared to the single-task problem, the presence of multiple tasks leads the planner to offer the farmer lower-powered incentives. Again, this is as our intuition would dictate. The planner realizing that runoff is now a strong risk complement will want to push the farmer to a less dispersed output vector. However, to ensure incentive compatibility for this less dispersed output vector, he must also offer the farmer a less dispersed payment vector.

To this point, we have assumed that p is a risk complement. However, it is easy to think of cases where p might be a risk substitute. Suppose that - indexes the presence or absence of pest infestations, and that p now measures runoff from the application of pesticides. It seems unlikely that pesticides will enhance crop output in the absence of a pest infestation so one might reasonably approximate this problem by assuming a cost structure of the form:

$$C(z) = c_1 z_1 + c_2 z_2$$

with $p(z) = \phi z_1$ with $\phi > 0$. That is increased runoff of pesticides is associated with higher production in the pest-infested state as a result of the application of increased pesticide application. Then our results change in a predictable fashion. Fixed payments and production standards are never optimal as this would encourage excessive marginal application of pesticides. And in comparison with the single-task case, the planner always offers the farmer a higher-powered incentive scheme. We leave the details of the analysis to the reader.

6 Concluding Remarks

We have studied the structure of state-contingent contracts in the presence of moral hazard and multi-tasking. Our analysis has identified necessary and sufficient conditions for the presence of multi-tasking to lead to fixed payments instead of incentive schemes. Holmström and Milgrom (1991) were the first to isolate the possibility of fixed payment schemes in the presence of moral hazard due to multi-tasking. We rely on the state-contingent representation of the moral hazard problem developed in our earlier work. That model places primary emphasis on the state-contingent technology, and consequently our analysis of the multi-tasking problem focuses on different aspects of the problem than the pathbreaking Holmström and Milgrom analysis. We show that when viewed from a state-contingent perspective, a primary determinant of whether multi-tasking leads to higher or lower powered incentive schemes is the role that the noncontractible outputs play in helping the agent deal with the production risk associated with the observable and contractible outputs. In particular, we show that when the noncontractible outputs are socially undesirable, standards are never optimal if they are also risk substitutes. If the noncontractible outputs are socially desirable, standards are never optimal if the noncontractible outputs play a risk-complementary role.

7 Appendix

Proof of Lemma 2: The incentive compatibility constraints in this case are:

$$\begin{aligned} \min\{y_1, y_2\} - C(z_1, z_2) &\geq y_1 - C(z_1, z_1) \\ \min\{y_1, y_2\} - C(z_1, z_2) &\geq y_2 - C(z_2, z_2) \\ C(z_2, z_1) - C(z_1, z_2) &\geq 0. \end{aligned}$$

Multiply the first constraint by λ $(0 < \lambda < 1)$ and the second constraint by $1 - \lambda$ and add the resulting equations together to obtain after rearrangement

$$\begin{aligned} \lambda C(z_1, z_1) + (1 - \lambda) C(z_2, z_2) - C(z_1, z_2) &\geq \lambda y_1 + (1 - \lambda) y_2 - \min\{y_1, y_2\} \\ &\geq 0. \end{aligned}$$

There are two possible ways to avoid a standard $z_2 > z_1$ and the reverse. Consider the former, then incentive compatibility requires

$$\lambda C(z_1, z_1) + (1 - \lambda) C(z_2, z_2) - C(z_1, z_2) \geq 0, \quad 0 < \lambda < 1.$$

Letting $\lambda \to 1$ yields contradicts the strict monotonicity of $C(z)$. A parallel argument yields a similar conclusion in the reverse case. This establishes necessity. Let $z_2 = z_1 = z$. Sufficiency is established by setting $y_1 = y_2 = \bar{u} + C(z, z)$.

Proof of Lemma 3: A multiplicative spread of z requires

$$\delta z_1 = -\frac{\pi_2}{\pi_1}\delta z_2.$$

with $\delta z_2 > 0$. Hence, by (15), any such change is mean preserving for y. The agent's payment in state 1 is

$$y_1 = h(\bar{u} + C(z_1, z_1)).$$

Observing that y_1 varies directly with z_1 then establishes the result.

Proof of Proposition 6: By (22) and (17) under the maintained assumption

$$h'\left[u_2^M\right] = \frac{\pi_2 - \mu\rho}{c_2} = h'\left[\bar{u} + c_2 z_2^{SB}\right].$$

Because h is strictly increasing and strictly convex, it follows immediately that

$$u_2^M = \bar{u} + c_2 z_2^{SB}.$$

This establishes that $y_2^M = y^{SB}$. The remainder of the inequality follows by Lemma 1 and the fact that under the maintained assumption a standard is not optimal. To prove the cost inequality, note that by the separability of the agent's objective function, the reservation utility constraint must bind exactly in both the second-best and moral hazard problems, i.e.,

$$\pi_1 u_1 + \pi_2 u_2 - C(z_1, z_2) = \bar{u}.$$

The first line of the theorem establishes that

$$\pi_1 u_1^M + \pi_2 u_2^M < u^{SB},$$

which with the reservation utility constraint yields cost inequality.

The last part is now obvious under strict monotonicity of the contract since we have established

$$c_2 z_2^{SB} > c_2 z_2^M + c_1 z_1^M.$$

Proof of Proposition 7: It follows trivially from the first-order conditions for the single-task and multiple task moral hazard problems that $u_2^S > u_2^M$, and therefore $y_2^S > y_2^M$. By (21) and (22)

$$\begin{aligned}
\pi_1 h'\left[u_1^M\right] &= \frac{1}{c_1 + c_2} - \pi_2 h'\left[u_2^M\right] - \frac{\mu\rho}{c_1 + c_2} \\
&= \frac{1}{c_1 + c_2} - \pi_2\left(\frac{\pi_2 - \mu\rho}{c_2} - \frac{\mu\rho}{c_1 + c_2}\right) \\
&= \frac{1}{c_1 + c_2} - \frac{\pi_2^2}{c_2} + \mu\rho\left(\frac{\pi_2}{c_2} - \frac{1}{c_1 + c_2}\right) \\
&> \frac{1}{c_1 + c_2} - \frac{\pi_2^2}{c_2} \\
&= \pi_1 h'\left[u_1^S\right],
\end{aligned}$$

where the inequality follows by (20). This establishes that $u_1^M > u_1^S$. The first series of inequalities now follows by Lemma 2 and the recognition that the contract must be monotonic under the current assumptions.

We first show that $z_1^M > z_1^S$. By (4)

$$u_1 = \underline{u} + (c_1 + c_2)\, z_1$$

in both the single-task and multi-task problems. Using the fact that $u_1^M > u_1^S$ then establishes that $z_1^M > z_1^S$. By (5)

$$u_2 = \underline{u} + (c_1 + c_2)\, z_1 + \frac{c_2}{\pi_2}\,(z_2 - z_1)$$

in both the single-task and multi-task problems. Because $u_2^S > u_2^M$, therefore,

$$(c_1 + c_2)\, z_1^S + \frac{c_2}{\pi_2} z_2^S > (c_1 + c_2)\, z_1^M + \frac{c_2}{\pi_2} z_2^M;$$

whence

$$\frac{c_2}{\pi_2}\left(z_2^S - z_2^M\right) > \left(c_1 + c_2 - \frac{c_2}{\pi_2}\right)\left(z_1^M - z_1^S\right):$$

By (20) and the fact that $z_1^M > z_1^1$, the right-hand side is positive, which establishes the result.