

Gaming and Strategic Ambiguity in Incentive Provision*

Florian Ederer[†]
MIT

Richard Holden[‡]
MIT and NBER

Margaret Meyer[§]
Oxford and CEPR

June 5, 2008

Abstract

A central tenet of economics is that people respond to incentives. While an appropriately crafted incentive scheme can achieve the second-best optimum in the presence of moral hazard, the principal must be very well informed about the environment (e.g. the agent's preferences and the production technology) in order to achieve this. Indeed it is often suggested that incentive schemes can be *gamed* by an agent with superior knowledge of the environment, and furthermore that lack of transparency about the nature of the incentive scheme can reduce gaming. We provide a formal theory of these phenomena. We show that random or ambiguous incentive schemes induce more balanced efforts from an agent who performs multiple tasks and who is better informed about the environment than the principal is. On the other hand, such random schemes impose more risk on the agent per unit of effort induced. By identifying settings in which random schemes are especially effective in inducing balanced efforts, we show that, if tasks are sufficiently complementary for the principal, random incentive schemes can dominate the best deterministic scheme. (*JEL* L13, L22)

*We thank Philippe Aghion, Patrick Bolton, Jeremy Bulow, Kate Doornik, Robert Gibbons, Edward Glaeser, Oliver Hart, Bengt Holmström, Ian Jewitt, Steven Matthews, Jesse Shapiro, Jeroen Swinkels, and Jean Tirole for helpful discussions and suggestions, as well as seminar participants at MIT.

[†]Department of Economics, E52-391, 50 Memorial Drive, Cambridge, MA, 02142, U.S.A., ederer@mit.edu.

[‡]MIT Sloan School of Management, E52-410, 50 Memorial Drive, Cambridge, MA, 02142, U.S.A., rholden@mit.edu.

[§]Nuffield College and Department of Economics, Oxford University, OX1 1NF, UK, margaret.meyer@nuffield.ox.ac.uk

1 Introduction

A fundamental consideration in designing incentive schemes is the possibility of *gaming*: the notion that an agent with superior knowledge of the environment to the principal can manipulate the incentive scheme to his own advantage. This is an important issue in theory as it suggests a reason why the second-best might not be attained and hence an additional source of efficiency loss. It is also an important practical matter. There is a large informal literature which documents the perverse effects of (high-powered) incentive schemes. This literature often concludes that unless the incentive designer can measure all relevant tasks extremely well (as is the objective of a “balanced scorecard”), she must inevitably trade off the negative effects of gaming against the positive ones from incentive provision.

It is also commonly suggested that lack of transparency—being deliberately ambiguous about the criteria which will be rewarded—can help circumvent gaming. This notion has a long intellectual history. It dates at least to Bentham (1830), who advocated the use of randomness in civil service selection tests.¹

One view as to why courts often prefer standards—which are somewhat vague—to specific rules is that it reduces incentives for gaming. For example, Weisbach (2000) argues that vagueness can reduce gaming of taxation rules, and Scott and Triantis (2006) argue that vague standards in contract law can improve *ex ante* incentives. Recently there have been calls for less transparency in the incentives provided to hospitals in the UK in the light of apparent gaming of incentive schemes that were designed to reduce patient waiting times (Bevan and Hood 2004). Similarly, the recent research assessment of UK universities has been marked by significant ambiguity about the criteria in an apparent attempt to deter gaming. There are numerous other examples in different, but related, incentive provision problems. The locations of speed cameras are often randomized,² security checks at airports and tax audits are often random, and even foreign policy often contains a significant degree of strategic ambiguity.

Despite the intuitive appeal of this line of argument, no formal theory has investigated it, and it is unclear how it relates to well-known economic theories of incentives. In the classic principal-agent model (Mirrlees 1974, Holmström 1979, Grossman and Hart 1983) the principal cannot observe the agent’s action(s), but knows his preferences, cost of effort, and the stochastic mapping from effort to output. The multi-task principal-agent model of Holmström and Milgrom (1991) gets closer to capturing the idea of gaming, by providing conditions under which incentives may optimally be very low-powered in response to the effort substitution problem. Yet there is still no role for ambiguity in this model. In some sense, the principal still knows “too much”.

In this paper we construct a formal theory of gaming and identify circumstances in which *ambiguity*, or lack of transparency, can be beneficial. Randomness is generally thought of as a bad

¹“Maximization of the inducement afforded to exertion on the part of learners, by impossibilizing the knowledge as to what part of the field of exercise the trial will be applied to, and thence making aptitude of equal necessity in relation to every part: thus, on the part of each, in so far as depends on exertion, maximizing the probable of absolute appropriate aptitude.” (Bentham, 1830/2005, Ch. IX, §16, Art 60.1)

²See Lazear (2006) for a model of this and other phenomena.

thing in moral hazard settings. Indeed, the central trade-off in principal-agent models is between insurance and incentives—and removing risk from the agent is desirable.³ Imposing less risk on the agent allows the principal to provide higher-powered incentives. In our model, however, randomness, despite having this familiar drawback, can nevertheless be beneficial overall, because it helps mitigate the undesirable consequences of the agent’s informational advantage.

In our model, the agent performs two tasks, which are substitutes in his cost-of-effort function, and receives compensation that is linear in his performance on each of the tasks, just as in Holmström and Milgrom (1991). The crucial difference is that there are two types of agent: one who has a lower cost of effort on task 1 and another who has a lower cost of effort on task 2. The principal’s benefit function is complementary in the efforts on the two tasks, so other things equal she prefers to induce both types of agent to choose balanced efforts, but the agent’s private information about his preferences makes this impossible to achieve with deterministic linear contracts. In this setting, it is advantageous to consider a richer contracting space, including random contracts. These random contracts make compensation ambiguous from the point of view of the agent, in that he knows that the compensation scheme ultimately used will take one of two possible forms (rewarding either task 1 or task 2 at a pre-specified rate) but, at the time he chooses his actions, he does not know which form will be used. Under *ex ante randomization*, the principal chooses randomly, before outputs are observed, which performance measure to reward. Under *ex post randomization*, the principal chooses which performance indicator to reward after observing outputs on the two tasks.

Our key contribution is to identify settings in which one or both of these randomized incentive schemes dominate all deterministic schemes. We identify three such environments. Each of these environments has the feature that one or both of the random contracts induce the agent to choose perfectly balanced efforts on the two tasks. The first such setting is that in which the agent has private information about his preferences but the magnitude of his preference across tasks is arbitrarily small. In this case, both *ex ante* and *ex post* randomization induce equal efforts on the two tasks from both types of agent. The second such setting is that where the shocks affecting outputs on the tasks are perfectly correlated. Here, *ex post* randomization induces perfectly balanced efforts. The final setting is the limiting case where the agents’ risk aversion becomes infinitely large and the variance of the shocks to outputs becomes arbitrarily small. In this case, both randomized schemes induce equal efforts on the two tasks. In all three environments, we show that there is a critical degree of complementarity of the principal’s benefit function above which the randomized contract(s) that induce(s) perfectly balanced efforts dominate(s) the best deterministic scheme.

Ex ante randomization pushes the agent toward balanced efforts on the tasks as a means of insuring himself against the risk generated by the random choice of which performance measure is rewarded. Under *ex post* randomization, there is an additional incentive to choose balanced efforts: The fact that the principal will choose to base compensation on the performance measure which minimizes her wage bill raises the agent’s expected marginal return to effort on the task on which

³For example, Holmström (1982) shows that in a multi-agent setting where agents’ outputs are correlated, the use of relative performance evaluation can remove risk from the agents and make it optimal to offer higher-powered incentive schemes.

he exerts less effort relative to the expected marginal return on the other task.

Our analysis shows that random contracts are more robust to uncertainty about the agent’s preferences than are deterministic ones. Specifically, we show that with random contracts, the ratio of the efforts exerted on the two tasks varies continuously when we introduce a small amount of uncertainty, whereas it varies discontinuously for deterministic contracts.

As the above discussion foreshadowed, our model can best be thought of in the light of two pathbreaking papers by Holmström and Milgrom (1987, 1991). In the first of these they provide conditions under which a linear contract is optimal. A key message of Holmström and Milgrom (1987) is that linear contracts are appealing because they are robust to the information structure.⁴ They illustrate this in the context of the Mirrlees (1974) result in which the first-best can be approximated by a highly non-linear incentive scheme. According to them, “to construct the [Mirrlees] scheme, the principal requires very precise knowledge about the agent’s preferences and beliefs, and about the technology he controls.” Holmström and Milgrom (1991) highlight that the *effort substitution problem* can lead to optimal incentives being extremely low-powered. When actions are technological substitutes for the agent, incentives on one task crowd out incentives on others.

There is also a large literature on subjective performance evaluation and relational contracts in which the principal has discretion over incentive payments (Bull 1987, MacLeod and Malcolmson 1989, Baker, Gibbons and Murphy 1994). As Prendergast (1999) points out, such discretion allows the principal “to take a more holistic view of performance; the agent can be rewarded for a particular activity only if that activity was warranted at the time”. Parts of this line of work share an important feature with our investigation: the agent (at least *ex ante*) has superior knowledge of the environment, e.g. Levin (2003). Unlike us, this literature focuses on how repeated interactions can allow for self-enforcing contracts, even when they are not verifiable by a court. In contrast, we show why even a *precise* contract rewarding multiple *verifiable* performance measures will often be problematic. We also explicitly model the risk imposed by introducing uncertainty about which performance measures will ultimately be used and show that despite this risk, a contract with randomization can dominate the best deterministic one. In models of subjective performance evaluation the (relational) contract itself imposes no additional risk on the agent, because there is common knowledge of equilibrium strategies by virtue of the solution concept employed for analyzing the repeated game.

An important paper by Bernheim and Whinston (1998) analyzes the incompleteness of observed contracts, a phenomenon they term “strategic ambiguity”. They show that when some aspects of an agent’s performance are non-contractible, it can be optimal not to specify other contingencies, even when these other contingencies are verifiable. Unlike us, they are focused on explaining optimal contractual incompleteness, rather than incentive provision in a moral hazard setting.

The paper perhaps most closely related to ours is MacDonald and Marx (2001). Like us, they analyze a principal-agent model with multiple tasks where the agent’s efforts on the tasks are substitutes in the agent’s cost function but complements in the principal’s benefit function, and

⁴A different strand of literature documents perverse incentives attributed to nonlinear schemes whereby agents make intertemporal effort shifts (eg. Asch (1990), Oyer (1998), among many others).

like us, they assume that the agent is privately informed about which of the two tasks he finds less costly. They, too, focus on how to design an incentive contract to overcome the agent's privately known bias and induce him to exert positive effort on both tasks. Since task outcomes are binary in their model, contracts consist of at most four distinct payments, and they show that the more complementary the tasks are for the principal, the more the agent's reward should be concentrated on the outcome where he produces two successes. While their model is designed to highlight the benefits of a simple type of nonlinear contract, they do not consider at all the benefits and costs of randomized incentive schemes.

Gjesdal's (1982) analysis of a single-task principal-agent model provides an example of a utility function for the agent for which randomization is beneficial.⁵ The benefit of randomization derives from the fact that the agent's first-order condition is convex in the payment from the principal for this utility function. Grossman and Hart (1983) show that the critical condition for ruling out randomness as optimal in such a setting is that the agent's preferences over income lotteries are independent of her action. A *sufficient* condition for this is that the agent's utility function is additively or multiplicatively separable in action and reward. This is stronger than required as it also implies that preferences for action lotteries are independent of income. The Sufficient Statistic Theorem of Holmström (1979) highlights this. There, the agent's utility function is additively separable, and his theorem implies that randomness cannot help because it does not affect the likelihood ratio. In our model, the agent has a multiplicatively separable utility function, and hence the attractiveness of random incentive schemes arises for quite different reasons than in Gjesdal (1982).

It is worth noting at this point that randomness and non-linearity are somewhat different concepts. One might argue that the linear contract used in Holmström and Milgrom (1991) is not optimal in a static setting and therefore that by adding additional features to the contract, it is hardly surprising that one can do better. To such an argument we have several responses. First, the special case of our model in which the principal is fully informed about the agent's preferences is precisely the Holmström-Milgrom setting, and we show that, in that special case, random contracts are *not* optimal. Thus, the attractiveness of random schemes arises because of the richer setting where the agent has superior knowledge of the environment. In fact, as we show, this superior knowledge can be arbitrarily small and still make random contracts optimal. Second, the random contracts in our model do not require the principal to commit to the randomizing procedure in advance. The outcome under ex ante randomization is equivalent to the equilibrium outcome of a game between the principal and the agent, and ex post randomization allows the principal to choose which performance measure to reward after outputs are realized. Therefore, our randomized schemes are feasible even when the principal is unable to commit to complicated non-linear contracts. *Indeed, we speculate that it may be the case that much of the appeal of random contracts is that they replicate complicated non-linear contracts in environments without commitment.*

The remainder of the paper is organized as follows. Section 2 outlines the model. In section

⁵The utility function is $U(s, a) = s(4 - a) - s^2/a$, where s is the payment and a is the action.

3 we introduce the classes of contracts we consider and analyze the equilibrium effort levels and profits under each. Section 4 identifies settings in which randomized contracts are dominated by one or more of the deterministic schemes. Section 5, which is the heart of the paper, identifies environments in which at least one of the random contracts can be shown to dominate the best deterministic scheme. Section 6 shows that our results are robust to various extensions such as relaxing the assumption that tasks are perfect substitutes for the agent and moving beyond the exponential-normal model. It also analyzes an additional type of random contract and identifies circumstances under which it is attractive. Section 7 offers some concluding remarks. Proofs not provided in the text are contained in the appendix.

2 The Model

A principal hires an agent to perform two tasks for her. The agent's output, x_j , on each task $j = 1, 2$ is observable and verifiable and depends both on the effort devoted by the agent to that task, e_j , and on the realization of a random shock, ε_j . Specifically, $x_j = e_j + \varepsilon_j$, where

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim N \begin{pmatrix} 0 & \sigma^2 & \rho\sigma^2 \\ 0 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

and ρ , the correlation between the shocks, is non-negative. The efforts chosen by the agent are not observable by the principal. In addition, the agent is privately informed about his costs of exerting efforts. With probability one-half, the agent's cost function is $c_1(e_1, e_2) = \frac{1}{2}(e_1 + \lambda e_2)^2$, in which case we will term him a type-1 agent, and with probability one-half his cost function is $c_2(e_1, e_2) = \frac{1}{2}(\lambda e_1 + e_2)^2$, in which case he will be termed a type-2 agent. We assume that the parameter $\lambda \geq 1$. For each type of agent $i = 1, 2$, efforts are perfect substitutes⁶: $\frac{\partial c_i / \partial e_1}{\partial c_i / \partial e_2}$ does not vary with (e_1, e_2) . Nevertheless, since $\lambda \geq 1$, each type of agent is biased towards a preferred task: for the type- i agent, the marginal cost of effort on task i is (weakly) lower than the marginal cost of effort on the other task. We assume that both types of agent have an exponential von Neumann-Morgenstern utility function with coefficient of absolute risk aversion r , so the type- i agent's utility function is

$$U = -e^{-r(w - c_i(e_1, e_2))},$$

where w is the payment from the principal. The two types of agent are assumed to have the same level of reservation utility, which we normalize to zero in certainty-equivalent terms.

An important feature of the model is that the agent's efforts on the tasks are complementary for the principal. We capture this by assuming that the principal's payoff is given by

$$\Pi = B(e_1, e_2) - w$$

⁶We relax this assumption in section 5—see also the supplementary material to this paper.

where the “benefit function” $B(e_1, e_2)$ takes the form

$$B(e_1, e_2) = \min\{e_1, e_2\} + \frac{1}{\delta} \max\{e_1, e_2\}.$$

The parameter $\delta \geq 1$ measures the degree of complementarity, with a larger value of δ implying greater complementarity. In the extreme case where $\delta = \infty$, the benefit function reduces to $B(e_1, e_2) = \min\{e_1, e_2\}$, and the efforts are perfect complements—this is the case where the principal’s desire for balanced efforts is strongest. At the other extreme, when $\delta = 1$, $B(e_1, e_2) = e_1 + e_2$, so the efforts are perfect substitutes—here the principal is indifferent as to how the agent allocates his total effort across the tasks.

The relative size of δ and λ determines what allocation of effort across tasks would maximize social surplus. If $\delta > \lambda$, so the principal’s desire for balanced efforts is stronger than the agent’s preference across tasks, then the surplus-maximizing effort allocation involves both types of agent exerting equal effort on the two tasks. If, instead, $\delta < \lambda$, then the first-best efficient effort allocation involves each type of agent focusing exclusively on his preferred task.

Throughout the analysis, we restrict attention to linear contracts of the form

$$w = \alpha + \beta_1 x_1 + \beta_2 x_2.$$

The distinction between deterministic and random contracts hinges on whether or not the values of α , β_1 , and β_2 are fully specified at the time the contract is signed. We say a contract is deterministic if, at the time the contract is signed, the agent is certain about what values of α , β_1 , and β_2 will be employed in determining his pay. If, instead, there is uncertainty about α , β_1 , and β_2 at the time of signing the contract, then we say the contract involves randomization.

3 Classes of Contracts

This section begins by studying deterministic contracts and shows how the form of the optimal deterministic scheme depends on the parameters of the environment, specifically the values of λ (which measures the strength of each type of agent’s preferences across tasks), δ (measuring the strength of the principal’s preference for balanced efforts), ρ (the correlation of the shocks affecting the outputs), and $R \equiv r\sigma^2$ (which represents the importance of risk aversion).⁷

We then introduce the two types of randomized contracts on which we focus. In the first, which we term *ex ante randomization*, the contract specifies that with probability p , $\beta_1 = \beta$ and $\beta_2 = 0$ and with probability $1-p$, $\beta_1 = 0$ and $\beta_2 = \beta$. Under *ex ante* randomization, the principal commits to employ a randomizing device to determine (with appropriate probabilities) on which of the two outputs the agent’s pay will be based. The second class of randomized contract is termed *ex post*

⁷For deterministic contracts, the values of r and σ^2 will affect the principal’s profits only through their product $r\sigma^2$, but as we will see below, for randomized contracts, r and σ^2 have separate influences on the agent’s effort choices and therefore on the principal’s profits.

randomization. Here, the principal, after observing the outputs x_1 and x_2 , chooses whether to pay $\alpha + \beta x_1$ or $\alpha + \beta x_2$. Under both types of randomization, the agent is ex ante uncertain about which performance indicator will determine his pay, but only under ex post randomization can the agent's efforts influence which indicator is ultimately used.⁸

3.1 Deterministic Contracts

3.1.1 The Special Case with Only One Type of Agent: $\lambda = 1$

We begin by analyzing the special case where $\lambda = 1$. There is only one type of agent, and since $c(e_1, e_2) = \frac{1}{2}(e_1 + e_2)^2$, he faces an equal marginal cost of effort on the two tasks. In this setting, the optimal deterministic contract can take one of two possible forms. The first form is a symmetric scheme, with $\beta_1 = \beta_2 = \beta$. Such a contract, which we denote by SD (for “symmetric deterministic”) can induce the agent to exert balanced efforts on the two tasks, but it exposes him to risk stemming from the random shocks affecting both tasks. The second form of contract rewards performance only on one task—it therefore induces the agent to exert effort only on one task, but exposes him to only one source of risk. This type of contract, which we denote by OT (for “one task”) has two mirror-image forms, and for concreteness we focus on the case in which task 1 is the task that is rewarded: $\beta_1 = \beta$ and $\beta_2 = 0$.⁹

Under the SD contract, $\beta_1 = \beta_2 = \beta$, and the agent is indifferent over all effort pairs that equate the common marginal cost of effort on the two tasks, $e_1 + e_2$, to the common marginal benefit, β . Since the parameter δ in the principal's benefit function is greater than or equal to one, the principal prefers the agent to choose $e_1 = e_2 = \frac{\beta}{2}$, and we assume that the agent does indeed choose this balanced effort allocation. The agent's certainty equivalent under the contract is

$$ACE = E(w) - c(e_1, e_2) - \frac{1}{2}r\sigma^2 Var(w) = \alpha + \beta^2 - \frac{\beta^2}{2} - R\beta^2(1 + \rho),$$

where $R \equiv r\sigma^2$. Given that the principal sets α to satisfy the agent's participation constraint with equality, the principal's expected profit as a function of β is

$$\Pi^{SD}(\beta, \lambda = 1) = \frac{\beta}{2} \left(1 + \frac{1}{\delta} \right) - \frac{\beta^2}{2} - R\beta^2(1 + \rho). \quad (1)$$

With β chosen optimally, the resulting maximized profit is

$$\Pi^{SD}(\lambda = 1) = \frac{(\delta + 1)^2}{8\delta^2 [1 + 2R(1 + \rho)]}. \quad (2)$$

Under the OT contract, $\beta_1 = \beta$ and $\beta_2 = 0$, so the agent sets $e_1 = \beta$ and $e_2 = 0$. With α chosen

⁸In Section 5 we will discuss another type of randomized contract, one in which the principal commits to determine pay according to $w = \alpha + \max\{\beta x_1, \beta x_2\}$.

⁹Given that $\rho \geq 0$, it *cannot* be optimal to set $\beta_1 > \beta_2 > 0$ or $\beta_2 > \beta_1 > 0$, since such contracts would expose the agent to risk stemming from both random shocks while failing to induce him to exert effort on more than one task.

optimally by the principal, the principal's expected profit as a function of β is

$$\Pi^{OT}(\beta, \lambda = 1) = \frac{\beta}{\delta} - \frac{\beta^2}{2} - \frac{1}{2}R\beta^2,$$

and the optimal choice of β yields profit

$$\Pi^{OT}(\lambda = 1) = \frac{1}{2\delta^2 [1 + R]}. \quad (3)$$

The SD contract induces the agent to exert effort on both tasks, while the OT contract elicits effort only on one task. However, since $\rho \geq 0$, for any given β the risk premium under the SD contract, $R\beta^2(1 + \rho)$, is larger than that under the OT contract, $\frac{1}{2}R\beta^2$. Therefore the principal faces a trade-off between the more balanced efforts induced by SD and the lower risk imposed by OT. Comparison of (2) and (3) shows that there is a critical value of the principal's complementarity parameter δ , greater than 1, above which SD is preferred and below which OT is preferred. This critical value is increasing in ρ and in R .

3.1.2 The General Case where the Agent has Private Information: $\lambda > 1$

In the general case where the agent is privately informed about his preferences across tasks ($\lambda > 1$), the optimal deterministic contract can take one of four possible forms.

The first form is the symmetric deterministic scheme introduced above, in which $\beta_1 = \beta_2 = \beta$. For any value of $\lambda > 1$, the SD scheme now induces both types of agent to focus their effort entirely on their preferred task: the type-1 agent chooses $e_1 = \beta$ and $e_2 = 0$, while the type-2 agent chooses $e_2 = \beta$ and $e_1 = 0$. The two types of agent receive equal expected utility from an SD contract, and with the fixed payment set optimally, the principal's expected profit as a function of β is

$$\Pi^{SD}(\beta, \lambda > 1) = \frac{\beta}{\delta} - \frac{\beta^2}{2} - R\beta^2(1 + \rho).$$

With β chosen optimally, the resulting maximized profit is

$$\Pi^{SD}(\lambda > 1) = \frac{1}{2\delta^2 [1 + 2R(1 + \rho)]}. \quad (4)$$

Note that as long as the principal's benefit function displays some degree of complementarity ($\delta > 1$), the profit the SD scheme drops discontinuously as λ is increased from 1, since for any $\lambda > 1$, both types of agent strictly prefer to focus their effort on one task, while at $\lambda = 1$, the agent is willing to choose the perfectly balanced effort allocation that the principal prefers.

The second candidate for an optimal deterministic scheme is a contract which rewards performance only on one task ($\beta_1 = \beta$ and $\beta_2 = 0$) and induces both types of agent to participate. We will continue to call this type of contract OT. Under the OT contract rewarding only task 1, the type-1 agent chooses $e_1 = \beta$ and $e_2 = 0$, while the type-2 agent chooses $e_1 = \frac{\beta}{\lambda}$ and $e_2 = 0$. The type-1 agent will derive strictly higher expected utility from this contract than the type-2 agent,

whose certainty equivalent is

$$ACE_2 = \alpha + \frac{\beta^2}{\lambda} - \frac{\beta^2}{2} - \frac{1}{2}R\beta^2.$$

With α chosen to satisfy the type-2 agent's participation constraint with equality, the principal's expected profit as a function of β is

$$\Pi^{OT}(\beta) = \frac{\beta}{2\delta} \left(\frac{\lambda+1}{\lambda} \right) - \frac{\beta^2}{2\lambda} (2\lambda - 1 + \lambda R),$$

so with β chosen optimally, maximized profit is

$$\Pi^{OT} = \frac{(\lambda+1)^2}{8\delta^2[(2\lambda-1)^2 + \lambda^2 R]}. \quad (5)$$

The third candidate for an optimal deterministic scheme shares with the OT contract the feature that it rewards performance only on one task, but it induces only one type of agent (the one for whom the rewarded task is the less costly one) to participate. Denote such a contract by OA (for “one agent”). The OA contract with $\beta_1 = \beta$ and $\beta_2 = 0$ induces the type-1 agent to choose $e_1 = \beta$ and $e_2 = 0$ and provides him with a certainty equivalent expected utility of

$$ACE_1 = \alpha + \frac{\beta^2}{2} - \frac{1}{2}R\beta^2.$$

With α chosen to satisfy the type-1 agent's participation constraint with equality, and thus to deter the type-2 agent from participating, the principal's expected profit as a function of β is

$$\Pi^{OA}(\beta) = \frac{\beta}{2\delta} - \frac{\beta^2}{4} - \frac{1}{4}R\beta^2,$$

The optimal choice of β yields profit

$$\Pi^{OA} = \frac{1}{4\delta^2[1+R]}. \quad (6)$$

When $\lambda > 1$, none of the three deterministic contracts analyzed above induces either type of agent to choose strictly positive effort on both tasks. There is in fact *no* deterministic contract of the form $w = \alpha + \beta_1 x_1 + \beta_2 x_2$ that can induce *both* types of agent to work on both tasks. Nevertheless, it is possible to design an asymmetric deterministic contract that offers positive but unequal rewards for the two tasks and that induces one type of agent to choose perfectly balanced efforts. This type of contract, which we denote by AD (for “asymmetric deterministic”), has two mirror-image forms, and for concreteness we focus on the case in which the type-2 agent is the one induced to exert balanced efforts, so task 1 is more highly rewarded: $\beta_1 = \lambda\beta_2$.

With an AD contract of this form, the type-2 agent is indifferent over all non-negative effort pairs (e_1, e_2) satisfying $\lambda e_1 + e_2 = \beta_2$. Among such effort pairs, the principal strictly prefers the type-2 agent to choose the perfectly balanced efforts $e_1 = e_2 = \frac{\beta}{1+\lambda}$ if and only if $\delta > \lambda$, i.e. if

and only if the principal's preference for balanced efforts outweighs the type-2 agent's preference for task 2. It is not hard to show that the best AD contract will be less profitable for the principal than the best SD contract whenever $\delta < \lambda$, so in the following we explicitly derive the profit from an AD contract only for the parameter region where $\delta \geq \lambda$.

Given the AD contract $\beta_1 = \lambda\beta_2$, the type-2 agent is willing to choose the perfectly balanced efforts preferred by the principal, $e_1 = e_2 = \frac{\beta}{1+\lambda}$, and he derives the certainty equivalent expected utility

$$ACE_2 = \alpha + \frac{1}{2}\beta_2^2 - \frac{1}{2}R(\lambda^2 + 2\lambda\rho + 1)\beta_2^2.$$

The type-1 agent chooses $e_1 = \lambda\beta_2$ and $e_2 = 0$, and his certainty equivalent under the contract is

$$\alpha + \frac{1}{2}\lambda^2\beta_2^2 - \frac{1}{2}R(\lambda^2 + 2\lambda\rho + 1)\beta_2^2.$$

Since $\lambda > 1$, this AD contract gives the type-2 agent strictly lower expected utility than the type-1 agent, so at the optimum, the participation constraint of the type-2 agent will be binding. With α set to achieve this, the principal's expected profit as a function of β_2 is

$$\Pi^{AD}(\beta_2) = \frac{1}{2} \frac{\beta_2(\lambda^2 + \lambda + \delta + 1)}{\delta(\lambda + 1)} - \frac{1}{2}\beta_2^2 [\lambda^2 + R(\lambda^2 + 2\lambda\rho + 1)]. \quad (7)$$

The optimal choice of β_2 yields profit

$$\Pi^{AD} = \frac{(\lambda^2 + \lambda + \delta + 1)^2}{8[\delta(\lambda + 1)]^2 [\lambda^2 + R(\lambda^2 + 2\lambda\rho + 1)]}. \quad (8)$$

The following arguments establish that there are no other deterministic contracts which can be more profitable than the best contract from among SD, OT, OA, and AD. First, consider a contract with $\beta_1 > \lambda\beta_2 > 0$. Such a contract would induce both types of agent to exert effort only on task 1. An alternative contract with the same value of β_1 but with β_2 set equal to zero would induce the same effort choices as the original contract but would impose less overall risk on the agents, given that $\rho \geq 0$. The principal can therefore earn greater profit from the alternative contract, because it allows him to set a lower fixed payment. This argument applies whether the original and alternative contract are designed to attract both types of agent (in which case the alternative is an OT scheme) or just the type-1 agent (in which case the alternative is an OA scheme).

Second, consider a contract with $\lambda\beta_2 > \beta_1 > \beta_2 > 0$. Such a contract would induce the type-1 agent to set $e_1 = \beta_1$ and $e_2 = 0$ and would induce the type-2 agent to set $e_1 = 0$ and $e_2 = \beta_2$. The optimal fixed payment would satisfy the type-2 agent's participation constraint with equality and leave rents to the type-1 agent. If the principal instead offered an SD contract with the common value of $\beta = \frac{1}{2}(\beta_1 + \beta_2)$, this alternative would yield the same expected benefit to the principal while allowing him to pay a lower expected wage (both because he would avoid leaving rents to the type-1 agent and because the risk premium would be lower). Hence the alternative SD contract would be more profitable for the principal.

Which of the four candidate deterministic contracts (SD, OT, OA, or AD) is most profitable for the principal will depend upon the parameters of the environment, specifically the values of λ (measuring the strength of each type of agent's preferences across tasks), δ (measuring the degree of complementarity of the tasks for the principal), ρ (the correlation of the shocks to outputs), and R (representing the importance of the agents' risk aversion). Since the main focus of our paper is on the properties and the profitability of randomized incentive schemes, it is not essential to provide an exhaustive comparison of the four deterministic schemes. Therefore, we simply provide a summary of the environments in which each of the four deterministic contracts dominates the other three. The summary below is illustrated in the three panels of Figure 1.

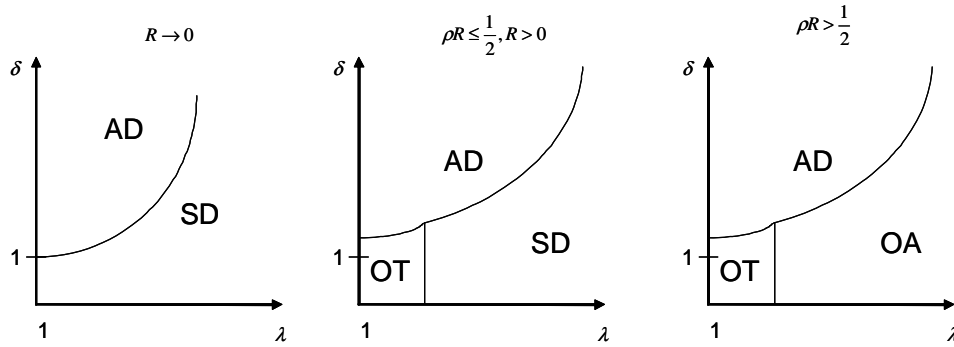


Figure 1: Deterministic Schemes

Since the AD contract induces one type of agent to choose balanced efforts, while the other three contracts induce fully focused efforts from both types, it is intuitive that if tasks are sufficiently complementary for the principal (δ sufficiently large), then the AD contract must be the preferred deterministic scheme. Furthermore, since the profit from the optimal AD contract falls as λ rises, while the profits from the optimal SD and OA contracts are independent of λ , it is also intuitive that the critical value of δ above which AD dominates SD and OA will be increasing in λ . The relative performance of SD and OA depends only on the size of ρR : since SD exposes the agent to risk stemming from the shocks to both tasks, while OA exposes him to risk from only one task, a larger value of ρR favors OA, and the critical value of ρR above which OA dominates SD is $\frac{1}{2}$. Since the profit from the optimal OT contract falls as λ rises (because it is increasingly costly, as well as less beneficial, to induce the second type of agent to participate), OT can be the optimal deterministic scheme only if both λ and δ are small. OT will in fact be optimal for small λ and δ unless we are in the limiting case where R goes to zero (so either the agents become risk neutral or the variance of the shocks to outputs goes to zero): as R goes to zero, the optimal deterministic contract is AD if δ is above some critical level and SD otherwise. Even as R goes to zero, this threshold value of δ exceeds λ (recall that λ is the critical value of δ above which first-best efficient efforts are balanced)—the reason is that AD, unlike SD, has the drawback that it gives rents to one type of agent.

3.2 Random Contracts

3.2.1 Ex Ante Randomization

A contract involving *ex ante randomization* specifies that with probability p , $\beta_1 = \beta$ and $\beta_2 = 0$ and with probability $1-p$, $\beta_1 = 0$ and $\beta_2 = \beta$. Under ex ante randomization, the principal commits to employ a randomizing device to determine (with appropriate probabilities) on which of the two outputs the agent's pay will be based.

Proposition 1 1. *Under ex ante randomization, it is optimal for the principal to commit to a value of p equal to $\frac{1}{2}$.*

2. *Let \bar{e}^{EAR} denote the effort exerted by each agent on his less-costly task and \underline{e}^{EAR} denote the effort exerted by each agent on his more-costly task. Then when ex ante randomization with $p = \frac{1}{2}$ induces interior solutions for the agents' effort choices, both types of agent choose $(\bar{e}^{EAR}, \underline{e}^{EAR})$ satisfying*

$$\bar{e}^{EAR} + \lambda \underline{e}^{EAR} = \frac{\beta}{\lambda + 1} \quad (9)$$

$$\lambda = \exp[r\beta(\bar{e}^{EAR} - \underline{e}^{EAR})]. \quad (10)$$

Thus

$$\begin{aligned} \bar{e}^{EAR} &= \frac{\beta}{(\lambda + 1)^2} + \frac{\lambda \ln \lambda}{r\beta(\lambda + 1)} \\ \underline{e}^{EAR} &= \frac{\beta}{(\lambda + 1)^2} - \frac{\ln \lambda}{r\beta(\lambda + 1)}, \end{aligned}$$

so the agents' effort choices will be interior solutions when $\beta^2 > \frac{(\lambda+1)\ln \lambda}{r}$.

3. *The principal's profit from interior effort choices by the agents under ex ante randomization with $p = \frac{1}{2}$, for a given β , is*

$$\Pi^{EAR}(\beta) = \underline{e}^{EAR} + \frac{1}{\delta} \bar{e}^{EAR} - \frac{\beta^2}{2(\lambda + 1)^2} - \frac{1}{2} R\beta^2 - \frac{\ln\left(\frac{(\lambda+1)^2}{4\lambda}\right)}{2r} \quad (11)$$

$$= \frac{(\delta + 1)\beta}{\delta(\lambda + 1)^2} - \frac{(\delta - \lambda)\ln \lambda}{\delta r\beta(\lambda + 1)} - \frac{\beta^2}{2(\lambda + 1)^2} - \frac{R\beta^2}{2} - \frac{\ln\left(\frac{(\lambda+1)^2}{4\lambda}\right)}{2r}. \quad (12)$$

Under ex ante randomization, it is optimal for the principal to commit to rewarding each of the two tasks with equal probability. Doing so results in the most balanced profile of effort choices (averaging across the two equally-likely types of agent), and also avoids leaving any rent to the type of agent whose less-costly task is more likely to be rewarded (if p is different from $\frac{1}{2}$).

To understand equations (9) and (10), note first that the sum of the expected marginal monetary returns to effort on the two tasks must be β , since either task 1 or task 2 is rewarded at rate β . If optimal efforts for the agents are interior, then adding the first-order conditions for effort on

the two tasks must yield $\beta = \partial c / \partial \bar{e} + \partial c / \partial \underline{e}$ for both types of agent, which gives us equation (9). Furthermore, for both types of agent, $\frac{\partial c / \partial \underline{e}}{\partial c / \partial \bar{e}} = \lambda$, and the first-order conditions for interior optimal efforts imply

$$\lambda = \frac{\frac{\partial c}{\partial \underline{e}}}{\frac{\partial c}{\partial \bar{e}}} = \frac{E[U'(\cdot)I_{\{\underline{x} \text{ is rewarded}\}}]}{E[U'(\cdot)I_{\{\bar{x} \text{ is rewarded}\}}]} = \frac{1-p}{p} \exp[r\beta(\bar{e} - \underline{e})], \quad (13)$$

which with $p = \frac{1}{2}$ gives us equation (10).

A contract involving ex ante randomization can induce the risk-averse agent to exert effort on both tasks (even when he finds one of them strictly less costly) as a means of partially insuring himself against the risk generated by the random choice of which task to reward. As equation (10) shows, optimal self-insurance will result in a smaller gap between the effort levels on the two tasks, the more risk-averse is the agent (larger r), the higher the incentive intensity (larger β), and the smaller the cost difference between tasks (smaller λ).

Recall that under a symmetric deterministic scheme, the agent's effort choices are discontinuous as λ is increased from one, switching from perfectly balanced efforts at $\lambda = 1$ (the allocation preferred by the principal) to completely focused efforts for any $\lambda > 1$. As a consequence, the principal's profit from a symmetric deterministic scheme drops discontinuously as λ is raised from one. In contrast, under ex ante randomization, both the agent's effort choices and the principal's profit are continuous in λ at $\lambda = 1$, as long as the agent is strictly risk-averse ($r > 0$). Thus we can say that ex ante randomization is more robust to the introduction of private information on the part of the agent than is a symmetric deterministic contract.¹⁰

Under ex ante randomization, the agent will ultimately be compensated based only on a single performance measure, but because (when $\lambda > 1$) the agent only partially insures himself against the risk associated with the random choice of measure, this scheme imposes greater risk costs on the agent than does a deterministic OT scheme which rewards only a single task at rate β . This can be seen in the principal's profit expression (11): whereas the risk premium under OT is $\frac{1}{2}R\beta^2$, the cost of the risk imposed on the agent under ex ante randomization is $\frac{1}{2}R\beta^2 + \frac{1}{2r} \ln\left(\frac{(\lambda+1)^2}{4\lambda}\right)$.

Remark 1 *We have established Proposition 1 under the assumption that the principal can commit to p . It is natural to wonder whether the same outcome would result if, instead, the principal chooses p at the same time as the agent chooses efforts (we term this “interim randomization”). Interim randomization induces a static game of incomplete information between the principal and the agent, and the appropriate solution concept is Bayes-Nash equilibrium. We can prove that under interim randomization, the unique Bayes-Nash equilibrium is the same as the outcome described in Proposition 1.¹¹ The attractive properties of ex ante randomization are thus not crucially dependent on the principal's having the power to commit to the randomizing probabilities.*

¹⁰Ex ante randomization is also more robust to uncertainty about the magnitude of λ than is an AD scheme. If the principal tries to design an AD scheme to induce one type of agent to choose balanced efforts but is even slightly wrong about the magnitude of λ , profit will be discontinuously lower than if he were right. The performance of ex ante randomization does not display this extreme sensitivity.

¹¹To see that the outcome described in Proposition 1 is an equilibrium under interim randomization, note that given that the two types of agent are equally likely and given that their effort choices are mirror images of each other, the principal anticipates equal expected output on the two tasks, so is willing to randomize over which one to

3.2.2 Ex Post Randomization

Under a contract involving *ex post randomization*, the principal, *after* observing the outputs x_1 and x_2 , chooses whether to pay $\alpha + \beta x_1$ or $\alpha + \beta x_2$. Just as under *ex ante* randomization, the agent is uncertain about which performance indicator will determine his pay, but with *ex post* randomization, unlike with *ex ante* randomization, the agent's choice of efforts can influence which indicator is ultimately used.

Since the principal will choose, *ex post*, to pay the smaller of the two possible wages, the agent anticipates that he will receive the wage

$$w = \min\{\alpha + \beta x_1, \alpha + \beta x_2\}.$$

To characterize the effort choices which maximize the agent's exponential expected utility, we make use of a result due to Cain (1994) which provides the moment-generating function for the minimum of bivariate normal random variables.

Proposition 2 1. *When ex post randomization induces interior solutions for the agents' effort choices, each type of agent chooses effort on his less-costly task, \bar{e}^{EPR} and effort on his more-costly task, \underline{e}^{EPR} satisfying*

$$\bar{e}^{EPR} + \lambda \underline{e}^{EPR} = \frac{\beta}{\lambda + 1} \quad (14)$$

$$\lambda = \exp[r\beta(\bar{e}^{EPR} - \underline{e}^{EPR})] \frac{\Phi\left(\frac{\bar{e}^{EPR} - \underline{e}^{EPR} + R\beta(1-\rho)}{\theta}\right)}{\Phi\left(\frac{-(\bar{e}^{EPR} - \underline{e}^{EPR}) + R\beta(1-\rho)}{\theta}\right)}, \quad (15)$$

where $\theta \equiv \sigma(1-\rho)^{\frac{1}{2}}$ and Φ is the c.d.f. of a standard normal random variable.

2. *The principal's profit from interior effort choices by the agents under ex post randomization, for a given β , is*

$$\begin{aligned} \Pi^{EPR}(\beta) &= \underline{e}^{EPR} + \frac{1}{\delta} \bar{e}^{EPR} - \frac{\beta^2}{2(\lambda + 1)^2} - \frac{1}{2} R\beta^2 \\ &\quad - \frac{1}{r} \ln \left[\frac{1 + \lambda}{\lambda} \Phi\left(\frac{\bar{e}^{EPR} - \underline{e}^{EPR} + R\beta(1-\rho)}{\theta}\right) \right] \\ &\quad - \beta(\bar{e}^{EPR} - \underline{e}^{EPR}) \Phi\left(-\frac{(\bar{e}^{EPR} - \underline{e}^{EPR})}{\theta}\right) + \beta\theta\phi\left(\frac{\bar{e}^{EPR} - \underline{e}^{EPR}}{\theta}\right), \end{aligned} \quad (16)$$

where ϕ is the density function of a standard normal random variable.

Proposition 2 shows that “aggregate” effort, $\bar{e} + \lambda \underline{e}$, is the same under *ex post* randomization as

reward. Given that the principal chooses $p = 1/2$ at equilibrium, the agents' optimal behavior is clearly as described in the proposition. To see that this outcome is the *unique* equilibrium, observe that if the agents conjectured that the principal would choose $p > 1/2$ ($p < 1/2$), then their optimal efforts would be such that the principal would anticipate larger expected output on task 1 (task 2), so the principal would strictly prefer to set $p = 0$ ($p = 1$).

under ex ante randomization—compare equations (14) and (9). Since both schemes reward either task 1 at rate β or task 2 at rate β , the sum of the expected marginal returns to effort on the two tasks must be β in both cases, and for interior solutions, this sum is equated to the sum of the marginal effort costs on the two tasks, $(\lambda + 1)(\bar{e} + \lambda \underline{e})$. Just as for ex ante randomization, the first-order conditions for interior optimal efforts imply

$$\lambda = \frac{\frac{\partial c}{\partial \underline{e}}}{\frac{\partial c}{\partial \bar{e}}} = \frac{E[U'(\cdot)I_{\{\underline{x} \text{ is rewarded}\}}]}{E[U'(\cdot)I_{\{\bar{x} \text{ is rewarded}\}}]}, \quad (17)$$

but for ex post randomization

$$\frac{E[U'(\cdot)I_{\{\underline{x} \text{ is rewarded}\}}]}{E[U'(\cdot)I_{\{\bar{x} \text{ is rewarded}\}}]} = \exp[r\beta(\bar{e} - \underline{e})] \frac{\Phi\left(\frac{\bar{e} - \underline{e} + R\beta(1-\rho)}{\theta}\right)}{\Phi\left(\frac{-(\bar{e} - \underline{e}) + R\beta(1-\rho)}{\theta}\right)},$$

which when combined with equation (17) gives us equation (15).

Under ex ante randomization, the risk-averse agent's incentive to choose (partially) balanced efforts derives purely from an insurance motive: a desire to insure himself against the risk generated by the random choice of which task to reward. Under ex post randomization, the insurance motive is still present, but because the principal will choose to base compensation on the task on which output is lower, there is an additional incentive for the agent to balance his efforts: as the gap $\bar{e} - \underline{e}$ between efforts on the less-costly and more-costly tasks widens, the likelihood that compensation will be based on output on the less-costly task falls, and this *per se* acts as a disincentive against raising $\bar{e} - \underline{e}$. Formally, the right-hand side of equation (15), which is increasing in $\bar{e} - \underline{e}$, is strictly greater than the right-hand side of equation (10) for all $\bar{e} - \underline{e} > 0$, so

$$\bar{e}^{EPR} - \underline{e}^{EPR} < \bar{e}^{EAR} - \underline{e}^{EAR} \quad \forall \lambda > 1. \quad (18)$$

Intuitively, we would expect that the principal's freedom, under ex post randomization, to choose the performance measure that minimizes his wage bill would result in weaker overall incentives for the agent than under ex ante randomization. This intuition is correct in the sense that the *sum* of the efforts on the two tasks, $\bar{e} + \underline{e}$, is lower under ex post than under ex ante randomization. However, efforts on the tasks are complementary in the principal's benefit function: $B(e_1, e_2) = \min\{e_1, e_2\} + \frac{1}{\delta} \max\{e_1, e_2\}$, where $\delta \geq 1$. It follows from equation (18) and the fact that $\bar{e} + \lambda \underline{e}$ is the same under the two schemes that the principal's expected benefit is higher or lower under ex post than under ex ante randomization according to whether δ , measuring the strength of his preference for balanced efforts, is larger or smaller than λ , measuring the strength of the agents' preference across tasks.

Equation (15) also allows us to show that under ex post randomization, the optimal gap between the effort levels on the two tasks is smaller the larger is r (because the stronger desire to self-insure is the dominant effect) and the smaller is λ (because it is less costly for the agent to choose balanced efforts). If $\lambda = 1$, both types of agent choose perfectly balanced efforts. These results parallel those

for ex ante randomization. Furthermore, while σ^2 and ρ have no effect on the gap in effort levels under ex ante randomization, under ex post randomization the effort gap is smaller the smaller is σ^2 and the larger is ρ . A smaller value of $\sigma^2(1 - \rho)$ makes any change in the agent's choice of $\bar{e} - \underline{e}$ more likely to affect which performance indicator is used, so gives the agent a stronger incentive to balance his efforts. For $\rho = 1$ (perfectly correlated shocks), optimal efforts are perfectly balanced. (We will study this case in detail below.)

We can also show that ex post randomization with coefficient β , in contrast to ex ante randomization, imposes *lower* risk costs on the agent than does a deterministic OT scheme which rewards only a single task at rate β , and therefore strictly lower risk costs than a SD scheme with coefficient β . Formally, this claim corresponds to the result that the principal's profit, given by equation (16), is *greater* than $\underline{e}^{EPR} + \frac{1}{\delta}\bar{e}^{EPR} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2}R\beta^2$. The intuitive reason why this result holds is that the variance of the wage under ex post randomization, $w = \min\{\alpha + \beta x_1, \alpha + \beta x_2\}$, is lower than the variance of either x_1 or x_2 by itself.¹²¹³

3.2.3 Ex Ante versus Ex Post Randomization

The preceding paragraphs have argued that i) ex post randomization induces a strictly smaller gap in efforts $\bar{e} - \underline{e}$ than ex ante randomization, while the two schemes induce the same aggregate effort $\bar{e} + \lambda \underline{e}$ and hence the same total cost of effort, and ii) ex post randomization imposes lower risk costs than ex ante randomization. Taken together, these findings generate the following proposition:

Proposition 3 *For any given $\beta > 0$, if $\delta \geq \lambda$, ex post randomization generates higher profit for the principal than ex ante randomization, and this comparison is strict as long as $\rho < 1$.*

The condition $\delta \geq \lambda$ ensures that the smaller gap in efforts under ex post randomization, coupled with the common value of $\max\{e_1, e_2\} + \lambda \min\{e_1, e_2\}$ under the two schemes, corresponds to a higher expected benefit for the principal. As long as $\rho < 1$, the risk costs imposed on the agent under ex post randomization are strictly lower than those imposed under ex ante randomization.

4 When Are Deterministic Contracts Optimal?

This section identifies three environments in which randomized contracts are dominated by one or more of the deterministic schemes. The first environment is that in which the agent has no private information about his preferences: $\lambda = 1$. The second is any setting where a randomized contract induces both types of agent to exert strictly positive effort on only one task. Finally, the third is that where $\delta \leq \lambda$, so the principal's preference for balanced efforts is weaker than the agent's preference

¹²We can also show that the risk costs imposed on the agent by ex post randomization are increasing in the gap $\bar{e} - \underline{e}$, reflecting the fact that the variance of $w = \min\{\alpha + \beta x_1, \alpha + \beta x_2\}$ is increasing in $\bar{e} - \underline{e}$.

¹³Note, however, that for a given coefficient β , the aggregate effort $\bar{e} + \lambda \underline{e}$ induced by ex post and ex ante randomization is only $\beta/(1 + \lambda)$, whereas under OT and SD it is β . Whether ex post randomization imposes greater or lower risk costs than OT and SD *per unit of aggregate effort induced* requires further analysis, and we pursue this question in Section 4.

across tasks. In each of these three environments, we can show that randomized contracts impose too much risk, relative to the effort benefits they generate, and as a consequence are dominated by a symmetric deterministic scheme.

Proposition 4 *When $\lambda = 1$, both ex ante and ex post randomization yield lower profit for the principal, for any given β , than a suitably designed symmetric deterministic scheme, and strictly lower profit for all $\rho < 1$.*

Proposition 4 shows that, when the principal is fully informed about the agent's preferences, so a deterministic contract is as effective at inducing balanced efforts as are the randomized schemes, then the randomized schemes are not attractive, because of the greater risk they impose on the agent per unit of effort induced. In order to explain this more fully, we now sketch the proof of Proposition 4.

As shown in section 3.1.1, when $\lambda = 1$, a symmetric deterministic contract with incentive coefficient β induces effort levels $\bar{e} = \underline{e} = \frac{\beta}{2}$ and hence yields the principal a profit of

$$\Pi^{SD}(\beta, \lambda = 1) = \frac{\beta}{2} \left(1 + \frac{1}{\delta}\right) - \frac{\beta^2}{2} - R\beta^2(1 + \rho).$$

From Proposition 1, ex ante randomization when $\lambda = 1$ induces effort levels $\bar{e} = \underline{e} = \frac{\beta}{4}$, and hence the principal's profit is

$$\Pi^{EAR}(\beta) = \frac{\beta}{4} \left(1 + \frac{1}{\delta}\right) - \frac{\beta^2}{8} - \frac{1}{2}R\beta^2.$$

From Proposition 2, ex post randomization when $\lambda = 1$ also induces effort levels $\bar{e} = \underline{e} = \frac{\beta}{4}$, and the principal's profit is

$$\Pi^{EPR}(\beta) = \frac{\beta}{4} \left(1 + \frac{1}{\delta}\right) - \frac{\beta^2}{8} - \frac{1}{2}R\beta^2 - \frac{1}{r} \ln \left[2\Phi \left(\frac{r\beta\sigma\sqrt{1-\rho}}{\sqrt{2}} \right) \right] + \frac{\beta\sigma\sqrt{1-\rho}}{\sqrt{\pi}}.$$

Now for a given value of β under ex post and ex ante randomization, set β under the symmetric deterministic scheme equal to $\beta^{SD} = \frac{\beta}{2}$. Then all three schemes induce the same effort levels and hence the same cost of effort. Therefore, any difference in profit between the SD scheme and the randomized schemes reflects entirely the differences in the costs of compensating the agent for the risk imposed on him. We have

$$\Pi^{SD} \left(\frac{\beta}{2} \right) - \Pi^{EAR}(\beta) = \frac{1}{2}R\beta^2 \left(1 - \frac{(1+\rho)}{2} \right), \quad (19)$$

and

$$\Pi^{SD} \left(\frac{\beta}{2} \right) - \Pi^{EPR}(\beta) = \frac{1}{2}R\beta^2 \left(1 - \frac{(1+\rho)}{2} \right) + \left\{ \frac{1}{r} \ln \left[2\Phi \left(\frac{r\beta\sigma\sqrt{1-\rho}}{\sqrt{2}} \right) \right] - \frac{\beta\sigma\sqrt{1-\rho}}{\sqrt{\pi}} \right\}. \quad (20)$$

It is clear by inspection that the right-hand side of (19) is strictly positive $\forall \beta > 0, \forall \rho < 1$, and hence ex ante randomization with incentive coefficient β is dominated by a symmetric deterministic

scheme with coefficient $\frac{\beta}{2}$. The additional term (in curly brackets) on the right-hand side of (20) is non-positive, and strictly negative for $\rho < 1$, and it reflects the lower risk costs imposed under ex post than under ex ante randomization (as shown in Proposition 3). The proof of Proposition 4 is completed by showing that, even though the term in curly brackets is negative, the right-hand side of (20) is strictly positive overall $\forall \beta > 0, \forall \rho < 1$. This implies that ex post randomization with incentive coefficient β is dominated by a symmetric deterministic scheme with coefficient $\frac{\beta}{2}$, which induces the same effort levels but imposes lower risk costs on the agent.

When $\lambda = 1$, we saw above that ex ante and ex post randomization require twice as large an incentive coefficient as a symmetric deterministic scheme to induce the same level of perfectly balanced efforts. For arbitrary λ , in order to induce any given *aggregate* effort level $\bar{e} + \lambda \underline{e}$, the two randomized schemes require an incentive coefficient that is $(\lambda + 1)$ times as large as that required by a symmetric deterministic scheme. Arguments very similar to those used to prove Proposition 4 establish, for arbitrary λ , that for any given level of aggregate effort $\bar{e} + \lambda \underline{e}$ induced by an ex ante or an ex post randomization scheme, a symmetric deterministic scheme can induce the same level of aggregate effort while imposing lower risk costs on the agent. (This last statement is equivalent to the special case of Proposition 6 below where $\delta = \lambda$.)

A second environment where randomized contracts are dominated by a deterministic scheme is any setting where a randomized contract induces both types of agent to exert strictly positive effort on only one task.

Proposition 5 *Whenever, for a given β , EAR or EPR induce the agent to exert effort only on his preferred task, the same effort allocation can be induced more profitably by a symmetric deterministic scheme.*

Intuitively, Proposition 5 holds for a similar reason to Proposition 4. Just as when $\lambda = 1$, when randomized contracts induce positive effort on only one task, they are providing no benefit relative to a deterministic scheme in the generation of balanced efforts, but they do have the cost of imposing greater risk on the agent for any given level of aggregate effort induced.

Proposition 5 has two informative corollaries:

Corollary 1 *Consider the limiting case where $\sigma^2 \rightarrow \infty$ and $r \rightarrow 0$ in such a way that $r\sigma^2 \rightarrow R \geq 0$. In this limiting case, for any $\beta > 0$, both ex ante randomization and ex post randomization induce the agent to exert effort only on his preferred task. Hence they are both dominated by a symmetric deterministic scheme.*

Recall that the profitability of deterministic schemes depends on the agent's risk aversion and the variance of the shocks to outputs only through the product $r\sigma^2$, whereas the performance of randomized contracts depends also on the individual values of r and σ^2 . Proposition 1 implies that as r falls, the gap in efforts under ex ante randomization rises (because the agent's desire for self-insurance diminishes), and Proposition 2 implies that as r falls and σ^2 rises, both changes lead to a larger gap in efforts under ex post randomization, both because the agent has less need to

self-insure and because larger σ^2 means that shifting effort from his less-preferred to his preferred task is more likely to raise the wage he ultimately receives. Corollary 1 shows that, for a given value of $r\sigma^2$, randomized schemes perform badly relative to deterministic ones when r is very small and σ^2 is very large, because in such settings, randomized schemes generate only very weak incentives to choose balanced efforts.

Corollary 2 *If*

$$J \equiv \left[1 + r\sigma^2(\lambda + 1)^2\right] \sqrt{\frac{(\lambda + 1)\ln \lambda}{r}} > 2, \quad (21)$$

then under ex ante randomization, when the incentive coefficient β is chosen optimally, the agent exerts effort only on his preferred task. Hence when condition (21) holds, ex ante randomization is dominated by a symmetric deterministic scheme.

Corollary 2 implies that, no matter how complementary the tasks are for the principal, if λ and/or σ^2 is sufficiently large, then ex ante randomization is less profitable than a symmetric deterministic scheme.

The final setting in which randomized contracts perform worse than deterministic ones is that in which $\delta \leq \lambda$:

Proposition 6 *When $\delta \leq \lambda$, then for any given β , both ex post randomization and ex ante randomization yield lower profit for the principal than a suitably designed symmetric deterministic scheme.*

If $\delta < \lambda$, it would not be efficient, even in the absence of moral hazard, for the principal to induce the agent to choose balanced efforts, because the benefit to the principal is outweighed by the disutility suffered by the agent. Hence in the presence of moral hazard (and private information about preferences), when $\delta < \lambda$, randomized contracts not only have the drawback of imposing more risk for any given level of aggregate effort induced, but they also induce inefficiently balanced efforts.

5 When Are Random Contracts Optimal?

We now investigate environments in which at least one of the random contracts can be shown to dominate the best deterministic scheme. We identify three such environments. Each of these settings has the feature that one or both of the random contracts induce the agent to choose perfectly balanced efforts. The first such setting is that in which the agent has private information about his preferences but the magnitude of his preference across tasks is arbitrarily small: this is the limiting case as $\lambda \rightarrow 1^+$. In this case, both ex ante and ex post randomization induce equal efforts on the two tasks from both types of agent. The second such setting is that where the shocks affecting outputs on the tasks are perfectly correlated: $\rho = 1$. Here, ex post randomization induces perfectly balanced efforts. The final setting is the limiting case where r goes to ∞ and σ^2 goes to 0 in such

a way that $r\sigma^2$ approaches $R \geq 0$. In this case, both ex ante and ex post randomization induce equal efforts on the two tasks. In all three environments, we show that there is a critical degree of complementarity of tasks for the principal above which the randomized contract(s) that induce(s) perfectly balanced efforts dominate(s) the best deterministic scheme.

5.1 The Limiting Case as $\lambda \rightarrow 1^+$

Under ex ante randomization, the agent's optimal efforts are continuous at $\lambda = 1$. In the limit as $\lambda \rightarrow 1$, $\bar{e} = \underline{e} = \frac{\beta}{4}$. As $\lambda \rightarrow 1$, the principal's profit as a function of β approaches

$$\Pi^{EAR}(\beta, \lambda = 1) = \frac{\beta}{4} \left(1 + \frac{1}{\delta}\right) - \frac{\beta^2}{8} - \frac{1}{2}R\beta^2.$$

With β chosen optimally, the resulting maximized profit is

$$\Pi^{EAR}(\lambda = 1) = \frac{(\delta + 1)^2}{8\delta^2 [1 + 4R]}. \quad (22)$$

By contrast, the agent's efforts and the principal's profits under the symmetric and asymmetric deterministic schemes are discontinuous at $\lambda = 1$. For $\lambda > 1$ but arbitrarily close to 1, which we term the limiting case as $\lambda \rightarrow 1^+$, Section 3.1.2 showed that the best deterministic contract is either the one-task scheme OT (if δ is small) or the asymmetric deterministic scheme AD (if δ is large)—see Figure 1. From equation (5), we have

$$\lim_{\lambda \rightarrow 1^+} \Pi^{OT} = \frac{1}{2\delta^2 (1 + R)}, \quad (23)$$

and from equation (8), we have

$$\lim_{\lambda \rightarrow 1^+} \Pi^{AD} = \frac{(3 + \delta)^2}{32\delta^2 [1 + 2R(1 + \rho)]}. \quad (24)$$

Comparison of (22) with (23) shows that for $\delta \leq \delta^{OT/EAR}$, $\Pi^{EAR} \leq \Pi^{OT}$, where

$$\delta^{OT/EAR} = 2 \left(\frac{1 + 4R}{1 + R} \right)^{\frac{1}{2}} - 1.$$

Note that $\delta^{OT/EAR}$ is increasing in R , going to 1 as $R \rightarrow 0$ and to 3 as $R \rightarrow \infty$. Comparison of (22) with (24) shows that for $\delta \leq \delta^{AD/EAR}$, $\Pi^{EAR} \leq \Pi^{AD}$, where

$$\delta^{AD/EAR} = \frac{3 - 2 \left(\frac{1 + 2(1 + \rho)R}{1 + 4R} \right)^{\frac{1}{2}}}{2 \left(\frac{1 + 2(1 + \rho)R}{1 + 4R} \right)^{\frac{1}{2}} - 1}.$$

$\delta^{AD/EAR}$ is increasing in R and decreasing in ρ , going to 1 as $\rho \rightarrow 1$ or $R \rightarrow 0$ and to $\frac{3 - \sqrt{2}}{\sqrt{2} - 1} = 3.82$

as $\rho \rightarrow 0$ and $R \rightarrow \infty$. We can thus conclude

Proposition 7 *Consider the limiting case as $\lambda \rightarrow 1^+$.*

1. *If δ , the degree of complementarity of the tasks in the principal's benefit function, is greater than*

$$\max \left\{ 2 \left(\frac{1+4R}{1+R} \right)^{\frac{1}{2}} - 1, \frac{3 - 2 \left(\frac{1+2(1+\rho)R}{1+4R} \right)^{\frac{1}{2}}}{2 \left(\frac{1+2(1+\rho)R}{1+4R} \right)^{\frac{1}{2}} - 1} \right\},$$

then ex ante randomization and ex post randomization both dominate the best deterministic scheme.

This critical value of δ is increasing in R and decreasing in ρ .

2. *For all $\rho \geq 0$ and $R \geq 0$, if $\delta > \frac{3-\sqrt{2}}{\sqrt{2}-1} = 3.82$, then ex ante and ex post randomization both dominate the best deterministic scheme.*

3. *As the shocks become perfectly correlated or $R \equiv r\sigma^2$ goes to 0, both randomized contracts dominate the best deterministic scheme for all $\delta > 1$.*

5.2 The Limiting Case of Perfect Correlation of the Shocks

Under ex post randomization, when the shocks to outputs are perfectly correlated, the agent, given his efforts on the two tasks, faces no uncertainty about which output (x_1 or x_2) compensation will be based on: He is certain that output will be lower, and hence compensation will be based, on the task on which he exerts lower effort. Therefore, starting from $e_1 = e_2 = e > 0$, the net return to the agent from increasing either e_1 alone or e_2 alone must be negative, because the marginal benefit is zero while the marginal cost is strictly positive. This implies that in searching for *either* type of agent's optimal (e_1, e_2) for a given β , we can confine attention to pairs (e_1, e_2) such that $e_1 = e_2 = e$.

For such pairs, the expected utility of both types of agent is

$$-\exp \left\{ -r \left[\alpha + \beta e - \frac{1}{2}(\lambda + 1)^2 e^2 - \frac{1}{2} R \beta^2 \right] \right\}$$

because the agent will receive a wage with the same distribution as $\alpha + \beta e + \beta u$, where $u \sim N(0, \sigma^2)$, that is, u has the same distribution as both shocks.

Therefore, both types of agent choose e according to the first order condition¹⁴

$$e = \frac{\beta}{(\lambda + 1)^2}.$$

With $\rho = 1$, the principal's profit given an arbitrary β and given that the fixed payment is set optimally is

$$\Pi^{EPR}(\beta) = \frac{\delta + 1}{\delta} \frac{\beta}{(\lambda + 1)^2} - \frac{\beta^2}{2(\lambda + 1)^2} - \frac{1}{2} R \beta^2. \quad (25)$$

¹⁴Formally, at $\rho = 1$ the agent's expected utility is not differentiable with respect to e_1 or e_2 at $e_1 = e_2$, but if we constrain $e_1 = e_2 = e$, expected utility is differentiable with respect to e .

Maximized profit is therefore

$$\Pi^{EPR} = \frac{(\delta + 1)^2}{2\delta^2(\lambda + 1)^2 [1 + R(\lambda + 1)^2]}. \quad (26)$$

The following lemma confirms that the principal's profit under ex post randomization when ρ is close to 1 is close to the profit level at $\rho = 1$.

Lemma 1 *Under ex post randomization, the agent's effort choices and the principal's profit are continuous at $\rho = 1$.*

Consider first the case where the principal's benefit function is the minimum of the efforts on the two tasks, i.e., the limiting case of $\delta \rightarrow \infty$.

In this case, for any $\lambda > 1$ and any R , the optimal deterministic scheme is AD, since this is the only deterministic scheme that does not induce both types of agent to choose corner solutions for efforts. Letting $\delta \rightarrow \infty$ in equation (7) gives us the profit under the AD scheme for any given β_2 (recall that the AD scheme sets $\beta_1 = \lambda\beta_2$):

$$\Pi^{AD}(\beta_2^{AD}) = \frac{\beta_2^{AD}}{2(\lambda + 1)} - \frac{(\beta_2^{AD})^2 \lambda^2}{2} - \frac{(\beta_2^{AD})^2}{2} R(\lambda + 1)^2.$$

We know from equation (53) that

$$\Pi^{EPR}(\beta^{EPR}) = \frac{\beta^{EPR}}{(\lambda + 1)^2} - \frac{(\beta^{EPR})^2}{2(\lambda + 1)^2} - \frac{(\beta^{EPR})^2}{2} R.$$

Setting $\beta^{EPR} = \frac{\beta_2^{AD}(\lambda + 1)}{2}$ gives

$$\Pi^{EPR}\left(\frac{\beta_2^{AD}(\lambda + 1)}{2}\right) = \frac{\beta_2^{AD}}{2(\lambda + 1)} - \frac{(\beta_2^{AD})^2}{8} - \frac{(\beta_2^{AD})^2}{8} R(\lambda + 1)^2,$$

and therefore the difference in profits is

$$\Pi^{EPR}\left(\frac{\beta_2^{AD}(\lambda + 1)}{2}\right) - \Pi^{AD}(\beta_2^{AD}) = \frac{(\beta_2^{AD})^2}{2} \left[\left(\lambda^2 - \frac{1}{4} \right) + \frac{3}{4} R(\lambda + 1)^2 \right] > 0,$$

$$\forall \beta_2^{AD} > 0, \forall \lambda > 1, \forall R \equiv r\sigma^2 \geq 0.$$

Thus, when the principal's preference for balanced efforts is extreme ($\delta \rightarrow \infty$), and the correlation of the shocks $\rho \rightarrow 1$, ex post randomization dominates the best deterministic scheme regardless of the values of λ and $R \equiv r\sigma^2$. The interpretation of this result is that when the principal's preference for balanced efforts is extreme, the fact that ex post randomization induces perfectly balanced efforts, regardless of the values of λ and R , outweighs the greater risk that ex post randomization imposes on the agent (per unit of aggregate effort induced).

Since the agent's behavior and the principal's profit under ex post randomization are analytically so tractable in the limiting case as the shocks become perfectly correlated, we can characterize

exactly when ex post randomization dominates all deterministic contracts and, if it does not, what deterministic contract is best. To keep the presentation relatively brief, we focus on three levels of the parameter R , which parallel the three panels of Figure 1. The following result is summarized by the three panels of Figure 2.

Proposition 8 *Consider the limit as $\rho \rightarrow 1$.*

1. *For $R \rightarrow 0$, for any given $\lambda > 1$, if δ is greater than (less than) λ , then the optimal contract is ex post randomization (a symmetric deterministic scheme).*
2. *For $R \in (0, 1/2]$, there exists $\hat{\lambda}(R) \in [1, 1.68]$ such that (i) if $\lambda \leq \hat{\lambda}(R)$, there is a critical δ above (below) which the optimal scheme is ex post randomization (a one-task deterministic contract) and (ii) if $\lambda > \hat{\lambda}(R)$, there is a critical δ above (below) which the optimal scheme is ex post randomization (a symmetric deterministic contract).*
3. *As $R \rightarrow \infty$, (i) if $\lambda \leq 1 + \sqrt{2}$, there exists a critical δ above (below) which the optimal scheme is ex post randomization (a one-task deterministic contract) and (ii) if $\lambda > 1 + \sqrt{2}$, there exist δ' and δ'' such that if $\delta > \delta'$, the optimal scheme is ex post randomization, if $\delta \in (\delta'', \delta')$, the optimal scheme is an asymmetric deterministic contract, and if $\delta < \delta''$, the optimal scheme is a deterministic contract that rewards only one task and attracts only one type of agent (OA).*

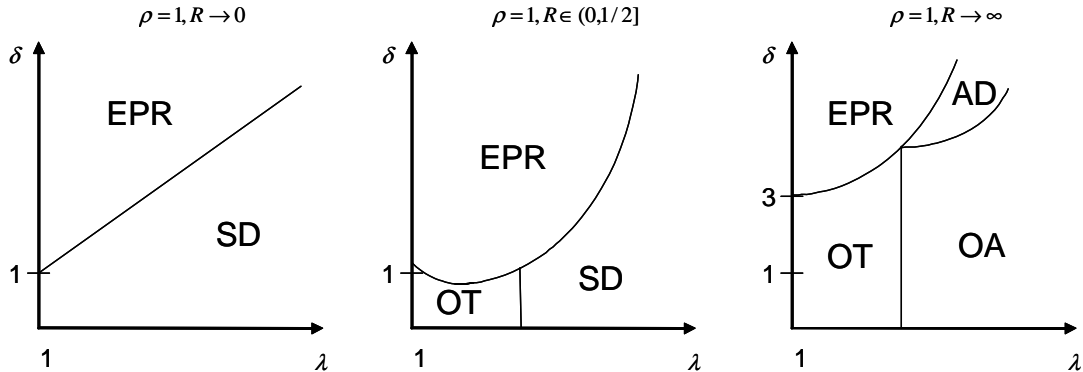


Figure 2: Optimality of Ex Post Randomization

The first part of the proposition shows that as R goes to zero, so the costs of imposing risk on the agent go to zero, the optimal incentive scheme is the one which induces the first-best effort allocation from the agents and which avoids paying any rent to one type of agent. For $\delta > \lambda$, ex post randomization achieves the first-best effort allocation, perfectly balanced efforts, when $\rho = 1$, and avoids paying any rent; for $\delta < \lambda$, the symmetric deterministic scheme achieves the first-best effort allocation, fully focused efforts, and avoids paying any rent. For R positive, the greater risk costs imposed by ex post randomization relative to the deterministic schemes must be traded off against the benefit of inducing perfectly balanced efforts from both agent types. For R not too large, the asymmetric deterministic (AD) scheme, which induces balanced efforts from one type of agent but pays the other type rents, can never be optimal; if δ is large, the best scheme is ex post

randomization, and if δ is small, the best scheme is one (OT or SD) which induces fully focused efforts. When R gets very large, so the greater risk costs imposed by ex post randomization become significant, then the critical δ above which ex post randomization is best rises and by so much that, for λ sufficiently large, there is an intermediate range of δ in which the AD scheme is optimal.

5.3 The Limiting Case where $r \rightarrow \infty$ and $\sigma^2 \rightarrow 0$

In Section 4, we considered the limiting case where $r \rightarrow 0$ and $\sigma^2 \rightarrow \infty$ in such a way that $r\sigma^2 \rightarrow R \geq 0$. We found that in this limit, both ex ante and ex post randomization induced the agent to exert effort only on his preferred task.

In the opposite limiting case where $r \rightarrow \infty$ and $\sigma^2 \rightarrow 0$ in such a way that $r\sigma^2 \rightarrow R \geq 0$, equation (10) in Proposition 1 shows that, for any $\beta > 0$, ex ante randomization induces the agent to choose perfectly balanced efforts: $\bar{e}^{EAR} - \underline{e}^{EAR} = 0$. This reflects the fact that as the agent becomes infinitely risk-averse, it becomes optimal to fully insure himself against the risk associated with the random choice of performance measure, by equalizing his expected outputs on the two tasks. Since comparison of equations (15) and (10) shows that $\bar{e}^{EPR} - \underline{e}^{EPR} \leq \bar{e}^{EAR} - \underline{e}^{EAR}$, it follows that in this limiting case, ex post randomization also induces perfectly balanced efforts—the reduction in σ^2 reinforces the agent's incentives for balance. Even if the product $r\sigma^2$ remains unchanged, so the profitability, as well as the efforts induced, under all deterministic schemes remains the same, when risk aversion becomes very large and exogenous shocks very small, both types of randomization generate very strong incentives to choose balanced efforts. As we show below, these very strong incentives for balance make both types of randomization perform better than any deterministic contract, as long as the principal's benefit function displays a sufficient degree of complementarity, δ .

Equations (11) and (16) in Propositions 1 and 2 show that, as $r \rightarrow \infty$ and $\sigma^2 \rightarrow 0$ in such a way that $r\sigma^2 \rightarrow R$, the principal's profit, under *either* type of randomization, approaches

$$\left(1 + \frac{1}{\delta}\right) \frac{\beta}{(\lambda + 1)^2} - \frac{\beta^2}{2(\lambda + 1)^2} - \frac{1}{2}R\beta^2.$$

Note that this is the same profit expression as under ex post randomization when $\rho \rightarrow 1$, but now ex ante randomization also achieves this profit level. Maximized profit is therefore

$$\Pi^{EAR} = \Pi^{EPR} = \frac{(\delta + 1)^2}{2\delta^2(\lambda + 1)^2 [1 + R(\lambda + 1)^2]}.$$

Profit under any deterministic scheme depends only on $r\sigma^2$, so in the limit we are considering, maximized profit under the asymmetric deterministic scheme is given by (see equation (8)):

$$\Pi^{AD} = \frac{(\lambda^2 + \lambda + \delta + 1)^2}{8[\delta(\lambda + 1)]^2 [\lambda^2 + R(\lambda^2 + 2\rho\lambda + 1)]}.$$

Comparing the previous two equations shows that the randomized schemes dominate AD if and

only if

$$\frac{(\delta + 1)^2}{(\lambda^2 + \lambda + \delta + 1)^2} \geq \frac{(1 + R(\lambda + 1))^2}{4[\lambda^2 + R(\lambda^2 + 2\rho\lambda + 1)]}.$$

We want to show that there is a critical δ above which the randomized schemes dominate AD. The left-hand side of the above inequality is increasing in δ , so it is sufficient to show that the inequality is satisfied as $\delta \rightarrow \infty$. The right-hand side will be largest for R as large as possible and ρ as small as possible. For $\delta \rightarrow \infty$, $R \rightarrow \infty$, and $\rho = 0$, the inequality reduces to

$$3\lambda^2 - 2\lambda + 3 \geq 0,$$

which is true for all $\lambda > 1$. Since for δ sufficiently large, the AD contract is the most profitable deterministic scheme (since it is the only deterministic contract that does not induce fully focused efforts from both types of agent), we have established the following:

Proposition 9 *Consider the limiting case where $r \rightarrow \infty$ and $\sigma^2 \rightarrow 0$ in such a way that $r\sigma^2 \rightarrow R \geq 0$.*

1. *Both ex ante and ex post randomization induce both types of agent to choose equal efforts on the two tasks, for any finite value of λ .*
2. *There is a critical value of δ above which both ex ante and ex post randomization are more profitable than any deterministic contract.*

6 Robustness and Extensions

6.1 Imperfect Substitutability of Efforts for the Agent

Throughout the analysis we have focused on the case where efforts are perfect substitutes in the agent's cost function. In the supplementary material to this paper, we extend our analysis to the case where efforts are imperfect substitutes. In particular, we study cost functions of the form

$$c(\bar{e}, \underline{e}) = \frac{1}{2} (\bar{e}^2 + 2s\lambda\bar{e}\underline{e} + \lambda^2\underline{e}^2) \quad (27)$$

where \bar{e} denotes each type of agent's effort on his preferred task and \underline{e} denotes each type's effort on the other task. In this parameterization, $s \in [0, 1]$ measures the degree of substitutability of efforts. Note that $s = 1$ represents perfect substitutability and $s = 0$ represents no substitutability.

With this more general cost function for the agent, a symmetric deterministic contract may now induce both types of agent to choose strictly positive effort levels on both tasks, even when the parameter λ is strictly greater than 1. In particular, if an agent's optimal efforts under a symmetric deterministic scheme with coefficient β on each task satisfy the first-order conditions $\beta = \frac{\partial c}{\partial \bar{e}} = \frac{\partial c}{\partial \underline{e}}$, then

$$\bar{e}^{SD} = \frac{\beta(\lambda^2 - s\lambda)}{\lambda^2(1 - s^2)} \text{ and } \underline{e}^{SD} = \frac{\beta(1 - s\lambda)}{\lambda^2(1 - s^2)}. \quad (28)$$

Thus a symmetric deterministic contract induces strictly positive efforts on both tasks from both types of agent if and only if $s\lambda < 1$, i.e. if and only if the substitutability of the tasks is significant enough relative to the strength of each agent's preference across tasks.

Since the analysis of this setting is rather voluminous, we direct the interested reader to the supplementary material. The key point to note is that our main finding is robust to the possibility of imperfect substitutability of efforts: It remains true that we can identify settings where the optimal deterministic scheme is dominated by a contract involving randomization.

6.2 Beyond The Exponential-Normal Model

Our findings that i) randomized incentive schemes induce more balanced efforts than deterministic schemes and that ii) they do so in a way that is robust to uncertainty about the agent's preferences apply even outside the exponential-normal model we have been considering. Suppose the production technology remains $x_1 = e_1 + \varepsilon_1$ and $x_2 = e_2 + \varepsilon_2$, but that $(\varepsilon_1, \varepsilon_2)$ have an arbitrary joint density $f(\varepsilon_1, \varepsilon_2)$ with identical marginal densities $g(\varepsilon_i)$. Suppose that each type of agent's utility is given by $U(w - c(\bar{e}, \underline{e}))$, where $c(\bar{e}, \underline{e})$ has the generalized form in (27), but we now let $U(\cdot)$ be an arbitrary strictly concave function.

Continue to consider ex ante and ex post randomization and focus on situations where each type of agent's optimal efforts are interior. Then for both types of randomization, it continues to follow from adding the first-order conditions for effort that

$$\beta = \frac{\partial c}{\partial \bar{e}} + \frac{\partial c}{\partial \underline{e}} \quad (29)$$

for each type of agent. It also follows that, for both types of randomization and for each type of agent,

$$\frac{\frac{\partial c}{\partial \underline{e}}}{\frac{\partial c}{\partial \bar{e}}} = \frac{E[U'(\cdot)I_{\{\underline{x} \text{ is rewarded}\}}]}{E[U'(\cdot)I_{\{\bar{x} \text{ is rewarded}\}}]}, \quad (30)$$

where \bar{x} is the output on the preferred task and \underline{x} is the output on the less-preferred task. For (symmetric) ex ante randomization, the right-hand side of (30) is

$$\frac{\int U'(\alpha + \beta(\underline{e} + \underline{u}) - c(\bar{e}, \underline{e}))g(\underline{u})d\underline{u}}{\int U'(\alpha + \beta(\bar{e} + \bar{u}) - c(\bar{e}, \underline{e}))g(\bar{u})d\bar{u}}. \quad (31)$$

For ex post randomization, the right-hand side of (30) is

$$\frac{\int \int U'(\alpha + \beta(\underline{e} + \underline{u}) - c(\bar{e}, \underline{e}))I_{\{\bar{e} + \bar{u} \geq \underline{e} + \underline{u}\}}f(\bar{u}, \underline{u})d\bar{u}d\underline{u}}{\int \int U'(\alpha + \beta(\bar{e} + \bar{u}) - c(\bar{e}, \underline{e}))I_{\{\bar{e} + \bar{u} < \underline{e} + \underline{u}\}}f(\bar{u}, \underline{u})d\bar{u}d\underline{u}}. \quad (32)$$

Under a symmetric deterministic scheme, an interior optimal solution for efforts still satisfies $\beta = \frac{\partial c}{\partial \bar{e}} = \frac{\partial c}{\partial \underline{e}}$, so just as in the exponential-normal model, an interior solution is given by (28) and exists if and only if $s\lambda < 1$.

Proposition 10 1. *When a symmetric deterministic scheme induces interior optimal efforts, ex ante randomization and ex post randomization do so as well. In this case ($s\lambda < 1$), both ex ante randomization and ex post randomization induce more balanced effort choices than the symmetric deterministic scheme, that is*

$$1 < \frac{\bar{e}^{EAR}}{\underline{e}^{EAR}} < \frac{\bar{e}^{SD}}{\underline{e}^{SD}}$$

and

$$1 < \frac{\bar{e}^{EPR}}{\underline{e}^{EPR}} < \frac{\bar{e}^{SD}}{\underline{e}^{SD}}.$$

2. *When efforts are perfect substitutes in the agent's cost function ($s = 1$) and we introduce a small amount of variation in the agent's preferences (λ is increased slightly from 1), then both $\frac{\bar{e}^{EAR}}{\underline{e}^{EAR}}$ and $\frac{\bar{e}^{EPR}}{\underline{e}^{EPR}}$ increase continuously from 1, whereas $\frac{\bar{e}^{SD}}{\underline{e}^{SD}}$ increases discontinuously (becoming infinite).*

Part 1 of the proposition confirms that even under these more general assumptions, both types of randomized incentive schemes induce more balanced efforts than deterministic schemes. Part 2 shows that they do so in a way that is more robust to uncertainty about the agent's preferences.

Just as in the exponential-normal model, under ex ante randomization the agent's incentive to choose more balanced efforts than under a symmetric deterministic scheme reflects purely an insurance motive: a desire to insure himself against the exogenous uncertainty about which indicator will be used to determine his compensation. Under ex post randomization, there is an additional force operating to induce balanced efforts: as the gap $\bar{e} - \underline{e}$ between efforts on the preferred and less preferred tasks widens, the likelihood that compensation will be based on output on the preferred task falls, and this *per se* acts as a disincentive against raising $\bar{e} - \underline{e}$.

Despite this additional incentive for balancing efforts, stronger assumptions would be needed to ensure in this more general setting that

$$\frac{\bar{e}^{EAR}}{\underline{e}^{EAR}} > \frac{\bar{e}^{EPR}}{\underline{e}^{EPR}}.$$

The reason is that marginal utility is evaluated over different sets of income realizations under ex ante and ex post randomization, so in contrast to the exponential-normal setting, the strength of the insurance motive is not generally equal under the two types of randomization.

In the limit as the agent becomes risk-neutral, however, a clear result is obtained since the insurance motive disappears. For ex ante randomization, (30) becomes

$$\frac{\frac{\partial C}{\partial \underline{e}}}{\frac{\partial C}{\partial \bar{e}}} = 1$$

and for ex post randomization, (30) becomes

$$\frac{\frac{\partial C}{\partial \underline{e}}}{\frac{\partial C}{\partial \bar{e}}} = \frac{\int \int I_{\{\bar{e} + \bar{u} \geq \underline{e} + \underline{u}\}} f(\bar{u}, \underline{u}) d\bar{u} d\underline{u}}{\int \int I_{\{\bar{e} + \bar{u} < \underline{e} + \underline{u}\}} f(\bar{u}, \underline{u}) d\bar{u} d\underline{u}}.$$

It follows that in the limit as the agent becomes risk-neutral,

$$1 < \frac{\bar{e}^{EPR}}{\underline{e}^{EPR}} \leq \frac{\bar{e}^{EAR}}{\underline{e}^{EAR}},$$

and the second inequality is certain to be strict whenever ex post randomization induces interior optimal efforts. Note that in this limit, ex ante randomization does no better than a symmetric deterministic scheme at inducing balanced efforts.

6.3 Another Random Contract: The “Max” Contract

In this section we consider the contract

$$w = \max \{ \alpha + \beta x_1, \alpha + \beta x_2 \}.$$

The first thing to note about this contract is that, unlike ex post randomization, this new contract requires the principal to be able to commit to the rule determining on which performance measure compensation will be based. Obviously, after outputs are realized, the principal would prefer to base compensation on the performance measure which minimizes her wage bill. Since the “max” contract pays according to the performance measure which *maximizes* her wage bill, commitment power is necessary for its implementation.

Nonetheless, under the assumption that the principal can commit, the contract exhibits a number of interesting features which we illustrate in the analysis below.

The key benefit of ex ante and ex post randomization is that they induce relatively balanced efforts from the agent under a wide range of parameter values. When tasks are strong enough complements for the principal, this benefit can be large enough to make these forms of randomization preferable to deterministic schemes. In contrast, the “max” contract always induces the agent to choose *unbalanced* efforts. It is intuitive, then, that the “max” contract can be attractive *only* when δ is small. We show below that there are settings in which the “max” contract is a more attractive way for the principal to induce fully focused efforts than is any deterministic scheme.

We first characterize the efforts induced by the “max” contract and the profit it generates for the principal. The derivation of these results closely parallels the derivations for ex post randomization, after adapting Cain’s (1994) methods to derive the moment-generating function for the *maximum* of bivariate normal random variables.

Proposition 11 1. *For all $\lambda > 1$, the “max” contract induces both types of agent to exert zero effort on their less-preferred task and effort \bar{e}^* on their preferred task, where \bar{e}^* satisfies the first-order condition*

$$\frac{\beta - \bar{e}}{\bar{e}} = \exp \{ r \beta \bar{e} \} \frac{\Phi \left(\frac{-\bar{e} - r \beta \sigma^2 (1 - \rho)}{\theta} \right)}{\Phi \left(\frac{\bar{e} - r \beta \sigma^2 (1 - \rho)}{\theta} \right)},$$

where $\theta \equiv \sigma(1 - \rho)^{\frac{1}{2}}$.

2. (i) $\bar{e}^* \in [\beta/2, \beta]$; (ii) \bar{e}^* is decreasing in r and decreasing in $\sigma\sqrt{1-\rho}$. (iii) $\lim_{\sigma\sqrt{1-\rho} \rightarrow \infty} \bar{e}^* = \beta/2$; (iv) \bar{e}^* is continuous at $\sigma\sqrt{1-\rho} = 0$, and $\lim_{\sigma\sqrt{1-\rho} \rightarrow 0} \bar{e}^* = \beta$.
3. The principal's profit under the “max” contract, for a given β is

$$\begin{aligned} \Pi^{\max}(\beta) = & \frac{\bar{e}^*}{\delta} - \frac{1}{2}(\bar{e}^*)^2 - \frac{1}{2}R\beta^2 - \frac{1}{r} \ln \left[\left(\frac{\beta}{\bar{e}^*} \right) \Phi \left(\frac{\bar{e}^* - R\beta(1-\rho)}{\theta} \right) \right] \\ & + \beta \bar{e} \Phi \left(\frac{-\bar{e}^*}{\theta} \right) - \beta \theta \phi \left(\frac{\bar{e}^*}{\theta} \right). \end{aligned} \quad (33)$$

Though part 1 of Proposition 11 is intuitive, the proof is not trivial, as there are two competing effects. On the one hand, the risk-averse agent has an incentive to insure himself against the risk associated with the uncertainty about which task compensation will be based on. This self-insurance motive *per se* pushes him towards balancing his efforts on the two tasks, just as it did under ex ante and ex post randomization. On the other hand, the fact that compensation under the “max” contract is based on the larger of the two task outputs *per se* gives the agent an incentive to focus all of his efforts on his preferred task. Proposition 11 shows that this second effect is the dominant one. This is proved by observing that, if there were an interior solution for efforts, then it would have to satisfy equation (30) adapted for the “max” contract:

$$\lambda = \frac{E[U'(\cdot)I_{\{\underline{x} \text{ is rewarded}\}}]}{E[U'(\cdot)I_{\{\bar{x} \text{ is rewarded}\}}]} = \exp\{r\beta(\bar{e} - \underline{e})\} \frac{\Phi\left(\frac{-(\bar{e}-\underline{e})-r\beta\sigma^2(1-\rho)}{\theta}\right)}{\Phi\left(\frac{(\bar{e}-\underline{e})-r\beta\sigma^2(1-\rho)}{\theta}\right)}. \quad (34)$$

It can be shown that the right-hand side of (34), evaluated at any $\bar{e} - \underline{e} \geq 0$, is less than or equal to 1. Since $\lambda > 1$, there can be no solution to (34) and hence no interior solution (\bar{e}, \underline{e}) for efforts under the “max” contract.

The dominant effect of an increase in risk aversion is the strengthening of the insurance effect, which pushes towards choosing a smaller gap in efforts, hence a smaller value of \bar{e} . On the other hand, an increase in $\sigma\sqrt{1-\rho}$ reduces the likelihood that an increase in effort on the preferred task will result in an increase in the wage, so lowers the optimal \bar{e} . Note that in contrast to ex post randomization, where a rise in risk aversion and a fall in $\sigma\sqrt{1-\rho}$ both induced a smaller gap in efforts, under the max contract these two changes have opposing effects on the effort gap and hence on the effort level \bar{e} . In the limiting case where $\sigma\sqrt{1-\rho} \rightarrow \infty$, the probability that the wage will be based on output on the preferred task falls to 1/2, so the optimal \bar{e} falls to $\beta/2$. In the other extreme, where $\sigma\sqrt{1-\rho} \rightarrow 0$, the agent is *certain* that any increase in \bar{e} by Δ (given $\underline{e} = 0$) will increase the wage by $\beta\Delta$, so the optimal effort on the preferred task rises to β , as high as under a symmetric deterministic scheme.

Given that the “max” contract induces both types of agent to exert zero effort on their less-preferred task, a necessary condition for it to be attractive is that tasks not be very complementary for the principal, i.e. that δ be small. But a small value δ is not sufficient, as shown by the following result.

Proposition 12 *Assume $\rho < 1$ and consider the limiting case where $r \rightarrow 0$ and $\sigma^2 \rightarrow \infty$ in such a way that $r\sigma^2 \rightarrow R \geq 0$. Then the “max” contract is dominated by the symmetric deterministic scheme.*

As Proposition 11, part 2, shows, in the limit as $\sigma^2 \rightarrow \infty$, the incentives provided by the “max” contract for effort on the preferred task are weakest, and this result is unaffected by $r \rightarrow 0$ at the same time. In the limiting case analyzed in Proposition 12, therefore, $\bar{e}^* \rightarrow \frac{\beta}{2}$, since regardless of the size of \bar{e} , the probability that increasing it will increase the wage received goes to 1/2. Ex post randomization, in this limiting case, also induces a corner solution for efforts (as shown in Corollary 1 in Section 4), and furthermore optimal effort on the preferred task under ex post randomization also approaches $\frac{\beta}{2}$ in this limit, for the same reason as under the “max” contract. We can also show that, in this limit, not only do these two contracts induce the same effort choices but also they generate the same profit for the principal, for any β . Therefore it follows from Corollary 1 that in this environment, the “max” contract is dominated by a symmetric deterministic scheme. Proposition 12 thus shows that when exogenous shocks are very large and risk aversion of the agent very small, the weak effort incentives provided by the max contract make it unattractive, even when inducing balanced efforts is not an important concern for the principal.

The most important contribution of this section is to identify conditions under which the “max” contract performs well. In particular, we can show that the max contract performs well when there is perfect correlation between the shocks to output, i.e. $\rho = 1$. As we explained above, when $\rho \rightarrow 1$, the marginal return to effort on the less-costly task (on which the agent focuses his effort) is highest, since the wage is certain to be based on performance on that task. The optimal effort level on that task is then as large as under a symmetric deterministic scheme or a one-task scheme that rewards the agent’s preferred task. Also, in general the “max” contract, for any given β , imposes weakly lower risk costs on the agent than a deterministic OT scheme with coefficient β , and therefore strictly lower risk costs than a SD scheme with coefficient β .¹⁵ The intuitive reason for this is the same as for the corresponding finding for ex post randomization: that the variance of the wage under the max contract, $w = \max\{\alpha + \beta x_1, \alpha + \beta x_2\}$ (like the variance of the wage under ex post randomization, $w = \min\{\alpha + \beta x_1, \alpha + \beta x_2\}$), is weakly lower than the variance of either x_1 or x_2 by itself. In the special case where $\rho = 1$, the max contract imposes exactly the same risk costs on the agent as the OT scheme.

Overall, therefore, the strong incentives for effort on the preferred task under the max contract when $\rho = 1$, coupled with its attractive risk properties, allow us to show that in this special case, the max contract is the most profitable way of inducing the agent to choose focused effort. Furthermore, in this case we can show that, regardless of the values of the other parameters in the model, every deterministic contract is always dominated by some form of randomized scheme, either one which bases the agent’s pay on the larger of his outputs (the “max” contract) or one which bases his

¹⁵Formally, this corresponds to the result that the principal’s profit, given by equation (33), is weakly greater than $\frac{\bar{e}^*}{\delta} - \frac{1}{2}(\bar{e}^*)^2 - \frac{1}{2}R\beta^2$.

pay on the smaller of his outputs (ex post randomization). These findings are summarized in the following proposition.

Proposition 13 *Suppose $\rho = 1$.*

- 1. Among the contracts which induce both types of agent to choose strictly positive effort on only one task, the “max” contract is the most profitable one for the principal, for all $\lambda > 1$, for all $\delta \geq 1$, and for all $R = r\sigma^2 \geq 0$.*
- 2. Every deterministic contract is dominated by either ex post randomization or the “max” contract, for all $\lambda > 1$, for all $\delta \geq 1$, and for all $R = r\sigma^2 \geq 0$.*

7 Conclusion

In this paper we have formalized the notion that an agent with superior knowledge of the contracting environment—here, his cost of effort on different tasks—may *game* an incentive scheme. Moreover, we have shown that random contracts—the principal being ambiguous in a well defined sense—can, in certain circumstances, alleviate this gaming. In such circumstances, the use of randomness helps redress the agent’s informational advantage by introducing uncertainty into the agent’s environment.

Our key contribution is to identify settings in which one or both of ex ante and ex post randomization dominate all deterministic incentive schemes. We identified three such environments. Each of these environments has the feature that one or both of the random contracts induce the agent to choose perfectly balanced efforts on the two tasks. The first such setting is that in which the agent has private information about his preferences but the magnitude of his preference across tasks is arbitrarily small. The second is that where the shocks affecting outputs on the tasks are perfectly correlated. The final setting is the limiting case where the agents’ risk aversion becomes infinitely large and the variance of the shocks to outputs becomes arbitrarily small. In all three of these environments, we showed that there is a critical degree of complementarity of tasks for the principal above which the randomized contract(s) that induce(s) perfectly balanced efforts dominate(s) the best deterministic scheme.

It is worth noting that the outcomes achieved under ex ante and ex post randomization in our model are achievable even if the principal cannot commit to a randomizing procedure in advance. The outcome under ex ante randomization is equivalent to the equilibrium outcome of a game between the principal and the agent, and ex post randomization allows the principal to choose which performance measure to reward after outputs are realized. Therefore, our randomized schemes are feasible even when the principal is unable to commit to complicated non-linear contracts. We suggest that part of the appeal of random contracts is that they replicate complicated non-linear contracts in environments without commitment.

We have taken a particular approach to modeling the agent’s superior knowledge of the environment. There are certainly other possibilities—such as the agent’s having private information about other components of her preferences than the cost of effort or about the stochastic mapping from effort to output. We have also restricted attention to a one-shot interaction. Future work could

analyze the benefits and costs of randomized incentive schemes in more general environments.

References

- Asch, Beth J.**, “Do Incentives Matter? The Case of Navy Recruiters,” *Industrial and Labor Relations Review*, Feb. 1990, *43* (3, Special Issue: Do Compensation Policies Matter?), 89S–106S.
- Baker, George, Robert Gibbons, and Kevin J. Murphy**, “Subjective Performance Measures in Optimal Incentive Contracts,” *Quarterly Journal of Economics*, Nov. 1994, *109* (4), 1125–1156.
- Bentham, Jeremy**, *Constitutional Code*, Vol. 1, London: R. Heward, 1830.
- Bernheim, B. Douglas and Michael D. Whinston**, “Incomplete Contracts and Strategic Ambiguity,” *American Economic Review*, 1998, *88*, 902–932.
- Bevan, Gwyn and Christopher Hood**, “Targets, Inspections, and Transparency,” *British Medical Journal*, May 2004, *328*, 598.
- Bull, Clive**, “The Existence of Self-Enforcing Implicit Contracts,” *Quarterly Journal of Economics*, Feb. 1987, *102* (1), 147–159.
- Cain, Michael**, “The Moment-Generating Function of the Minimum of Bivariate Normal Random Variables,” *The American Statistician*, May 1994, *48* (2), 124–125.
- Gjesdal, Froystein**, “Information and Incentives: The Agency Information Problem,” *Review of Economic Studies*, Jul. 1982, *49* (3), 373–390.
- Grossman, Sanford J. and Oliver D. Hart**, “An Analysis of the Principal-Agent Problem,” *Econometrica*, Jan. 1983, *51* (1), 7–45.
- Holmstrom, Bengt**, “Moral Hazard and Observability,” *Bell Journal of Economics*, Spring 1979, *10* (1), 74–91.
- , “Moral Hazard in Teams,” *Bell Journal of Economics*, Autumn 1982, *13* (2), 324–340.
- and **Paul Milgrom**, “Aggregation and Linearity in the Provision of Intertemporal Incentives,” *Econometrica*, Mar. 1987, *55* (2), 303–328.
- and —, “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design,” *Journal of Law, Economics, and Organization*, 1991, *7* (Special Issue: Papers from the Conference on the New Science of Organization, January 1991), 24–52.
- Lazear, Edward P.**, “Speeding, Terrorism, and Teaching to the Test,” *Quarterly Journal of Economics*, Aug. 2006, *121* (3), 1029–1061.
- Levin, Jonathan**, “Relational Incentive Contracts,” *American Economic Review*, Jun. 2003, *93* (3), 835–857.

- MacDonald, Glenn and Leslie M. Marx**, “Adverse Specialization,” *Journal of Political Economy*, Aug. 2001, *109* (4), 864–899.
- MacLeod, W. Bentley and James M. Malcomson**, “Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment,” *Econometrica*, Mar. 1989, *57* (2), 447–480.
- Mirrlees, James A.**, “Notes on Welfare Economics, Information and Uncertainty,” in M. Balch, D. McFadden, and S. Wu, eds., *Essays in Equilibrium Behavior under Uncertainty*, Amsterdam: North-Holland, 1974.
- Oyer, Paul**, “Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality,” *Quarterly Journal of Economics*, Feb. 1998, *113* (1), 149–185.
- Prendergast, Canice**, “The Provision of Incentives in Firms,” *Journal of Economic Literature*, Mar. 1999, *37* (1), 7–63.
- Scott, Robert E. and George G. Triantis**, “Anticipating Litigation in Contract Design,” *Yale Law Journal*, Jan. 2006, *115* (4), 814–879.
- Weisbach, David A.**, “An Efficiency Analysis of Line Drawing in the Tax Law,” *Journal of Legal Studies*, Jan. 2000, *29* (1), 71–97.

8 Appendix

8.1 Omitted Proofs

Proof of Proposition 1. Agent 1 maximizes expected utility

$$\begin{aligned} E[-\exp(-r(w - c(e)))] &= -p \exp\left(-r\left(\alpha + \beta e_1 - \frac{r}{2}\sigma^2\beta^2 - c_1(e)\right)\right) \\ &\quad - (1-p) \exp\left(-r\left(\alpha + \beta e_2 - \frac{r}{2}\sigma^2\beta^2 - c_1(e)\right)\right) \end{aligned}$$

The first-order conditions are

$$\begin{aligned} p[\beta - (e_1 + \lambda e_2)] \Delta_1^1 - (1-p)(e_1 + \lambda e_2) \Delta_2^1 &= 0 \\ -p(e_1 + \lambda e_2) \lambda \Delta_1^1 + (1-p)[\beta - (e_1 + \lambda e_2) \lambda] \Delta_2^1 &= 0. \end{aligned}$$

where

$$\begin{aligned} \Delta_1^1 &= \exp\left(-r\left(\alpha + \beta e_1 - \frac{r}{2}\sigma^2\beta^2 - c_1(e)\right)\right) \\ \Delta_2^1 &= \exp\left(-r\left(\alpha + \beta e_2 - \frac{r}{2}\sigma^2\beta^2 - c_1(e)\right)\right) \end{aligned}$$

Adding the first-order conditions gives

$$p[\beta - (e_1 + \lambda e_2)(\lambda + 1)] \Delta_1^1 + (1-p)[\beta - (e_1 + \lambda e_2)(\lambda + 1)] \Delta_2^1 = 0$$

\Leftrightarrow

$$[\beta - (e_1 + \lambda e_2)(\lambda + 1)] [p\Delta_1^1 + (1-p)\Delta_2^1] = 0$$

so that we have

$$e_1 + \lambda e_2 = \frac{\beta}{(\lambda + 1)}.$$

Now substituting this into either one of the first-order conditions and rearranging yields

$$p\lambda\Delta_1^1 = (1-p)\Delta_2^1.$$

\Leftrightarrow

$$\ln \frac{p\lambda}{1-p} = r\beta(e_1 - e_2).$$

Solving the system of equations, we find for agent 1

$$\begin{aligned} e_1 &= \frac{\beta}{(\lambda + 1)^2} + \frac{\lambda \ln \frac{p\lambda}{1-p}}{r\beta(\lambda + 1)} \\ e_2 &= \frac{\beta}{(\lambda + 1)^2} - \frac{\ln \frac{p\lambda}{1-p}}{r\beta(\lambda + 1)}. \end{aligned}$$

For agent 2, analogous steps yield optimal effort levels

$$\begin{aligned}
e_1 &= \frac{\beta}{(\lambda+1)^2} - \frac{\ln \frac{(1-p)\lambda}{p}}{r\beta(\lambda+1)} \\
e_2 &= \frac{\beta}{(\lambda+1)^2} + \frac{\lambda \ln \frac{(1-p)\lambda}{p}}{r\beta(\lambda+1)}.
\end{aligned}$$

The maximized expected utilities are

$$\begin{aligned}
EU_1 &= -p(\lambda+1) \exp \left(-r \left(\alpha + \frac{\beta^2}{2(\lambda+1)^2} - \frac{r\sigma^2\beta^2}{2} + \frac{\lambda \ln \frac{p\lambda}{1-p}}{r(\lambda+1)} \right) \right) \\
&= -A_1 \exp \left(-r \left(\alpha + \frac{\beta^2}{2(\lambda+1)^2} - \frac{r\sigma^2\beta^2}{2} \right) \right) \\
EU_2 &= -(1-p)(\lambda+1) \exp \left(-r \left(\alpha + \frac{\beta^2}{2(\lambda+1)^2} - \frac{r\sigma^2\beta^2}{2} + \frac{\lambda \ln \frac{(1-p)\lambda}{p}}{r(\lambda+1)} \right) \right) \\
&= -A_2 \exp \left(-r \left(\alpha + \frac{\beta^2}{2(\lambda+1)^2} - \frac{r\sigma^2\beta^2}{2} \right) \right)
\end{aligned}$$

where

$$\begin{aligned}
A_1 &= p(\lambda+1) \exp \left(-\frac{\lambda \ln \frac{p\lambda}{1-p}}{\lambda+1} \right) \\
A_2 &= (1-p)(\lambda+1) \exp \left(-\frac{\lambda \ln \frac{(1-p)\lambda}{p}}{\lambda+1} \right).
\end{aligned}$$

Comparing these expressions we can see that $EU_1 < EU_2$ when $p < \frac{1}{2}$ and $EU_1 > EU_2$ when $p > \frac{1}{2}$. Hence at the optimum the IR constraint will be binding for agent 1 when $p < \frac{1}{2}$ and for agent 2 when $p > \frac{1}{2}$. Note that the problem is entirely symmetric around $p = \frac{1}{2}$, so we need only focus on $p < \frac{1}{2}$.

We now use agent 1's binding IR constraint to find α

$$p(\lambda+1) \exp \left[-r \left(\alpha + \frac{\beta^2}{2(\lambda+1)^2} - \frac{r\sigma^2\beta^2}{2} + \frac{\lambda \ln \frac{p\lambda}{1-p}}{r(\lambda+1)} \right) \right] = 1$$

\Leftrightarrow

$$\alpha = -\frac{\beta^2}{2(\lambda+1)^2} + \frac{r\sigma^2\beta^2}{2} - \frac{\lambda \ln \frac{p\lambda}{1-p}}{r(\lambda+1)} + \frac{\ln(p(\lambda+1))}{r}$$

Now denote agent i 's effort on task j as e_j^i . Then the principal's expected wage bill is

$$\begin{aligned}
E[w] &= \alpha + \frac{p}{2}\beta e_1^1 + \frac{1-p}{2}\beta e_2^1 + \frac{p}{2}\beta e_1^2 + \frac{1-p}{2}\beta e_2^2 \\
&= -\frac{\beta^2}{2(\lambda+1)^2} + \frac{r\sigma^2\beta^2}{2} - \frac{\lambda \ln \frac{p\lambda}{1-p}}{r(\lambda+1)} + \frac{\ln(p(\lambda+1))}{r} \\
&\quad + \frac{p}{2} \left[\frac{\beta^2}{(\lambda+1)^2} + \frac{\lambda \ln \frac{p\lambda}{1-p}}{r(\lambda+1)} \right] + \frac{1-p}{2} \left[\frac{\beta^2}{(\lambda+1)^2} - \frac{\ln \frac{p\lambda}{1-p}}{r(\lambda+1)} \right] \\
&\quad + \frac{p}{2} \left[\frac{\beta^2}{(\lambda+1)^2} - \frac{\ln \frac{(1-p)\lambda}{p}}{r(\lambda+1)} \right] + \frac{1-p}{2} \left[\frac{\beta^2}{(\lambda+1)^2} + \frac{\lambda \ln \frac{(1-p)\lambda}{p}}{r(\lambda+1)} \right] \\
&= \frac{\beta^2}{2(\lambda+1)^2} + \frac{r\sigma^2\beta^2}{2} + K(p, \lambda, r)
\end{aligned}$$

where

$$K(p, \lambda, r) = -\frac{\lambda \ln \left(\frac{p\lambda}{1-p} \right)}{r(\lambda+1)} + \frac{\ln(p(\lambda+1))}{r} + \frac{[p(\lambda+1) - 1] \ln \left(\frac{p\lambda}{1-p} \right)}{2r(\lambda+1)} + \frac{[\lambda - p(\lambda+1)] \ln \left(\frac{(1-p)\lambda}{p} \right)}{2r(\lambda+1)}$$

The principal's profit function is

$$\Pi^{EAR} = \frac{1}{2} \left[e_2^1 + \frac{1}{\delta} e_1^1 + e_1^2 + \frac{1}{\delta} e_2^2 \right] - E[w],$$

and substituting for $E[w]$ yields

$$\begin{aligned}
\Pi^{EAR} &= \frac{1}{2} \left\{ \frac{\beta}{(\lambda+1)^2} - \frac{\ln \frac{p\lambda}{1-p}}{r\beta(\lambda+1)} + \frac{1}{\delta} \left[\frac{\beta}{(\lambda+1)^2} + \frac{\lambda \ln \frac{p\lambda}{1-p}}{r\beta(\lambda+1)} \right] \right\} \\
&\quad + \frac{1}{2} \left\{ \frac{\beta}{(\lambda+1)^2} - \frac{\ln \frac{(1-p)\lambda}{p}}{r\beta(\lambda+1)} + \frac{1}{\delta} \left[\frac{\beta}{(\lambda+1)^2} + \frac{\lambda \ln \frac{(1-p)\lambda}{p}}{r\beta(\lambda+1)} \right] \right\} - E[w] \\
&= \frac{(\delta+1)\beta}{\delta(\lambda+1)^2} - \frac{(\delta-\lambda) \ln \lambda}{\delta r\beta(\lambda+1)} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{r\sigma^2\beta^2}{2} + K(p, \lambda, r).
\end{aligned}$$

In this expression for profit, we have assumed that $\frac{1}{2} > p > \frac{1}{\lambda+1}$. The profit can be shown to be even lower when $p < \frac{1}{\lambda+1}$. Note also that since $K(p, \lambda, r)$ does not depend on β , the optimal choice of β is independent of the mixing probability p . Furthermore, $K(p, \lambda, r)$ is increasing in p for $p < \frac{1}{2}$ since

$$K(p, \lambda, r) = -\frac{\lambda \ln \left(\frac{p\lambda}{1-p} \right)}{r(\lambda+1)} + \frac{\ln(p(\lambda+1))}{r} + \frac{[p(\lambda+1) - 1] \ln \left(\frac{p\lambda}{1-p} \right)}{2r(\lambda+1)} + \frac{[\lambda - p(\lambda+1)] \ln \left(\frac{(1-p)\lambda}{p} \right)}{2r(\lambda+1)}$$

and hence

$$\frac{\partial K}{\partial p} = -\frac{\lambda-1}{2rp(1-p)(\lambda+1)} + \frac{\ln \frac{p}{1-p}}{r} < 0.$$

Thus the optimal choice of p is $p^* = \frac{1}{2}$.

When $p = \frac{1}{2}$ we have for agent 1

$$\begin{aligned} e_1 &= \frac{\beta}{(\lambda+1)^2} + \frac{\lambda \ln \lambda}{r\beta(\lambda+1)} \\ e_2 &= \frac{\beta}{(\lambda+1)^2} - \frac{\ln \lambda}{r\beta(\lambda+1)} \end{aligned}$$

so that $e_1 > e_2$. Similarly for agent 2

$$\begin{aligned} e_1 &= \frac{\beta}{(\lambda+1)^2} - \frac{\ln \lambda}{r\beta(\lambda+1)} \\ e_2 &= \frac{\beta}{(\lambda+1)^2} + \frac{\lambda \ln \lambda}{r\beta(\lambda+1)} \end{aligned}$$

so that $e_2 > e_1$. With $p = \frac{1}{2}$, the optimal effort level on the preferred task is the same for each agent, as is the optimal effort level on the less preferred task. Denoting the former by \bar{e} and the latter by \underline{e} , we have

$$\begin{aligned} \bar{e} + \lambda \underline{e} &= \frac{\beta}{(\lambda+1)} \\ \bar{e} - \underline{e} &= \frac{\ln \lambda}{r\beta}. \end{aligned}$$

These efforts will constitute interior solutions to the first-order conditions when $\underline{e} > 0$, i.e. when $\beta^2 > \frac{(\lambda+1)\ln \lambda}{r}$.

With $p = \frac{1}{2}$, the agents' maximized expected utilities are equal, so neither type of agent earns rents. The optimal value of α is

$$\alpha = -\frac{1}{r} \ln \left(\frac{2}{\lambda+1} \right) - \frac{\lambda \ln \lambda}{r(\lambda+1)} - \frac{\beta^2}{2(\lambda+1)^2} + \frac{r\sigma^2\beta^2}{2}.$$

The expected wage payment is therefore given by

$$E[w] = \frac{\ln \left(\frac{(\lambda+1)^2}{4\lambda} \right)}{2r} + \frac{\beta^2}{2(\lambda+1)^2} + \frac{r}{2}\sigma^2\beta^2,$$

so

$$\Pi^{EAR} = \frac{(\delta+1)\beta}{\delta(\lambda+1)^2} - \frac{(\delta-\lambda)\ln \lambda}{\delta r\beta(\lambda+1)} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{r\sigma^2\beta^2}{2} - \frac{\ln \left(\frac{(\lambda+1)^2}{4\lambda} \right)}{2r}.$$

■

Proof of Proposition 2. Since the agent's expected utility depends on $E \exp(-r\beta \min\{x_1, x_2\})$, we use the moment generating function for the minimum of bivariate normal random variables:

$$\begin{aligned} m(t) = \exp(t\mu_1 + \frac{1}{2}t^2\sigma_1^2) \Phi \left(\frac{\mu_2 - \mu_1 - t(\sigma_1^2 - \rho\sigma_1\sigma_2)}{\theta} \right) \\ + \exp(t\mu_2 + \frac{1}{2}t^2\sigma_2^2) \Phi \left(\frac{\mu_1 - \mu_2 - t(\sigma_2^2 - \rho\sigma_1\sigma_2)}{\theta} \right) \end{aligned}$$

where Φ is the c.d.f. of a standard normal random variable, μ_1 and μ_2 are the means of x_1 and x_2 , and $\theta \equiv (\sigma_2^2 - 2\rho\sigma_1\sigma_2 + \sigma_1^2)^{\frac{1}{2}}$. Since the principal's expected wage depends on $E \min\{x_1, x_2\}$, we use the formula (derived from the moment-generating function):

$$E \min\{x_{i1}, x_{i2}\} = \mu_1 \Phi\left(\frac{\mu_2 - \mu_1}{\theta}\right) + \mu_2 \Phi\left(\frac{\mu_1 - \mu_2}{\theta}\right) - \theta \phi\left(\frac{\mu_2 - \mu_1}{\theta}\right)$$

where ϕ is the density function of a standard normal random variable. For more details see Cain (1994).

An agent of type 1 chooses his effort levels to maximize the following expression

$$\begin{aligned} U_1 &= -\exp\left(-r\alpha + \frac{r}{2}(e_1 + \lambda e_2)^2\right) E[\exp(-r\beta \min\{x_1, x_2\})] \\ &= -\exp\left(-r\alpha + \frac{r}{2}(e_1 + \lambda e_2)^2\right) m(-r\beta) \end{aligned}$$

where m is the moment generating function of $\min\{x_1, x_2\}$.

The first order condition with respect to e_1 is

$$\begin{aligned} 0 &= -r(e_1 + \lambda e_2)m(-r\beta) \\ &\quad + r\beta \exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right) \Phi\left(\frac{e_2 - e_1 + r\beta\sigma^2(1-\rho)}{\theta}\right) \\ &\quad + \frac{1}{\theta} \exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right) \phi\left(\frac{e_2 - e_1 + r\beta\sigma^2(1-\rho)}{\theta}\right) \\ &\quad - \frac{1}{\theta} \exp\left(-r\beta e_2 + \frac{1}{2}r^2\beta^2\sigma^2\right) \phi\left(\frac{e_1 - e_2 + r\beta\sigma^2(1-\rho)}{\theta}\right). \end{aligned} \quad (35)$$

Similarly for e_2 we have

$$\begin{aligned} 0 &= -\lambda r(e_1 + \lambda e_2)m(-r\beta) \\ &\quad + r\beta \exp\left(-r\beta e_2 + \frac{1}{2}r^2\beta^2\sigma^2\right) \Phi\left(\frac{e_1 - e_2 + r\beta\sigma^2(1-\rho)}{\theta}\right) \\ &\quad + \frac{1}{\theta} \exp\left(-r\beta e_2 + \frac{1}{2}r^2\beta^2\sigma^2\right) \phi\left(\frac{e_1 - e_2 + r\beta\sigma^2(1-\rho)}{\theta}\right) \\ &\quad - \frac{1}{\theta} \exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right) \phi\left(\frac{e_2 - e_1 + r\beta\sigma^2(1-\rho)}{\theta}\right). \end{aligned} \quad (36)$$

Adding the two first order conditions we find

$$e_1 + \lambda e_2 = \frac{\beta}{(\lambda + 1)}. \quad (37)$$

Expanding the third and fourth terms in the two first-order conditions (35) and (36) reveals that in both FOC's these terms net to 0 for all (e_1, e_2) , and hence for (35) we have

$$(e_1 + \lambda e_2)m(-r\beta) = \beta \exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right) \Phi\left(\frac{e_2 - e_1 + r\beta\sigma^2(1-\rho)}{\theta}\right).$$

Substituting into this using (37) yields

$$m(-r\beta) = (\lambda + 1) \exp \left(-r\beta e_1 + \frac{1}{2} r^2 \beta^2 \sigma^2 \right) \Phi \left(\frac{e_2 - e_1 + r\beta \sigma^2 (1 - \rho)}{\theta} \right)$$

\Leftrightarrow

$$\lambda = \exp [r\beta(e_1 - e_2)] \frac{\Phi \left(\frac{e_1 - e_2 + r\beta \sigma^2 (1 - \rho)}{\theta} \right)}{\Phi \left(\frac{e_2 - e_1 + r\beta \sigma^2 (1 - \rho)}{\theta} \right)}. \quad (38)$$

Both factors on the RHS of (38) are increasing in $e_1 - e_2$. As a result, the optimal value of $e_1 - e_2$ is increasing in λ . If $\lambda = 1$, the optimal value of $e_1 - e_2 = 0$. Straightforward differentiation shows that the RHS of (38) is increasing in ρ for $e_1 - e_2 > 0$, so the optimal value of $e_1 - e_2$ is decreasing in ρ (if $\lambda > 1$).

Since ex post randomization treats the two tasks symmetrically ex ante, and since the two types of agent are mirror images of each other, the type-2 agent's optimal efforts on his preferred and less-preferred tasks will match the optimal values for the type-1 agent when this labeling is used.

Denote the level of effort each type chooses on his preferred task by \bar{e}^{EPR} and on his less-preferred task by \underline{e}^{EPR} . Define $d^{EPR} \equiv \bar{e}^{EPR} - \underline{e}^{EPR}$.

Using (37) and (38), we can express the maximized expected utility of both types under ex post randomization as

$$U = -\exp \left\{ -r \left[\alpha - \frac{\beta^2}{2(\lambda + 1)^2} - \frac{1}{2} r \sigma^2 \beta^2 \right] \right\} \frac{1 + \lambda}{\lambda} \times \exp (-r\beta \underline{e}^{EPR}) \Phi \left(\frac{\bar{e}^{EPR} - \underline{e}^{EPR} + r\beta \sigma^2 (1 - \rho)}{\theta} \right)$$

\Leftrightarrow

$$U = -\exp \left\{ -r \left[\left(\alpha + \beta \underline{e}^{EPR} - \frac{\beta^2}{2(\lambda + 1)^2} - \frac{1}{2} r \sigma^2 \beta^2 \right) \right] \right\} \times \exp \left\{ -r \left[-\frac{1}{r} \ln \left[\frac{1 + \lambda}{\lambda} \Phi \left(\frac{\bar{e}^{EPR} - \underline{e}^{EPR} + r\beta \sigma^2 (1 - \rho)}{\theta} \right) \right] \right] \right\}.$$

For both types of agent, the certainty equivalent is

$$ACE = \alpha + \beta \underline{e}^{EPR} - \frac{\beta^2}{2(\lambda + 1)^2} - \frac{1}{2} r \sigma^2 \beta^2 - \frac{1}{r} \ln \left[\frac{1 + \lambda}{\lambda} \Phi \left(\frac{\bar{e}^{EPR} - \underline{e}^{EPR} + r\beta \sigma^2 (1 - \rho)}{\theta} \right) \right] \quad (39)$$

while the principal's expected profit is

$$\begin{aligned} \Pi^{EPR} &= \underline{e}^{EPR} + \frac{1}{\delta} \bar{e}^{EPR} - \alpha - \beta E \min\{x_1, x_2\} \\ &= \underline{e}^{EPR} + \frac{1}{\delta} \bar{e}^{EPR} - \alpha - \beta \left[\bar{e}^{EPR} \Phi \left(-\frac{d^{EPR}}{\theta} \right) + \underline{e}^{EPR} \Phi \left(\frac{d^{EPR}}{\theta} \right) - \theta \phi \left(\frac{d^{EPR}}{\theta} \right) \right] \\ &= \underline{e}^{EPR} + \frac{1}{\delta} \bar{e}^{EPR} - \alpha - \beta \underline{e}^{EPR} - \beta d^{EPR} \Phi \left(-\frac{d^{EPR}}{\theta} \right) + \beta \theta \phi \left(\frac{d^{EPR}}{\theta} \right). \end{aligned} \quad (40)$$

Using (39) to substitute into (40) yields the principal's expected profit

$$\begin{aligned}\Pi^{EPR} = & \underline{e}^{EPR} + \frac{1}{\delta} \bar{e}^{EPR} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2} r \sigma^2 \beta^2 \\ & - \frac{1}{r} \ln \left[\frac{1+\lambda}{\lambda} \Phi \left(\frac{\bar{e}^{EPR} - \underline{e}^{EPR} + r\beta\sigma^2(1-\rho)}{\theta} \right) \right] - \beta d^{EPR} \Phi \left(-\frac{d^{EPR}}{\theta} \right) + \beta \theta \phi \left(\frac{d^{EPR}}{\theta} \right).\end{aligned}$$

■

Proof of Proposition 3. Under ex ante randomization, the principal's expected profit can be expressed as

$$\Pi^{EAR} = \underline{e}^{EAR} + \frac{1}{\delta} \bar{e}^{EAR} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2} r \sigma^2 \beta^2 - \frac{1}{r} \ln \frac{1+\lambda}{2\lambda} - \frac{\beta}{2} d^{EAR}.$$

The difference between the two profit expressions is

$$\begin{aligned}\Pi^{EPR} - \Pi^{EAR} = & \left[\left(\underline{e}^{EPR} + \frac{1}{\delta} \bar{e}^{EPR} \right) - \left(\underline{e}^{EAR} + \frac{1}{\delta} \bar{e}^{EAR} \right) \right] \\ & + \left\{ -\frac{1}{r} \ln \left[2\Phi \left(\frac{d^{EPR} + r\beta\sigma^2(1-\rho)}{\theta} \right) \right] - \beta \left[d^{EPR} \Phi \left(-\frac{d^{EPR}}{\theta} \right) - \frac{d^{EAR}}{2} \right] + \beta \theta \phi \left(\frac{d^{EPR}}{\theta} \right) \right\}.\end{aligned}\quad (41)$$

The first term (in square brackets) in (41) is the difference in the principal's benefits between ex post and ex ante randomization. The second term (in curly brackets) is the difference in expected costs of compensating the agent. Note that since both schemes induce the same aggregate effort $\bar{e} + \lambda \underline{e}$, there is no difference between them in the costs of effort incurred. The difference in costs of compensating the agent stems entirely from the difference in risk imposed on the agent by the two schemes.

It is straightforward to show that, given equations (18), (9), and (14), the first term (in square brackets) in (41) is non-negative if $\delta \geq \lambda$.

We now show that the second term (in curly brackets) in (41) is non-negative for all $\lambda \geq 1$, $\forall \beta > 0$, and strictly positive as long as $\rho < 1$. This finding reflects the fact that ex post randomization exposes the agent to less risk than ex ante randomization, even when both schemes induce unequal efforts on the two tasks.

Using (14) and (41), we can express d^{EAR} in terms of d^{EPR}

$$\begin{aligned}\exp(r\beta d^{EAR}) &= \exp[r\beta d^{EPR}] \frac{\Phi \left(\frac{d^{EPR} + r\beta\sigma^2(1-\rho)}{\theta} \right)}{\Phi \left(\frac{-d^{EPR} + r\beta\sigma^2(1-\rho)}{\theta} \right)} \\ \Leftrightarrow r\beta d^{EAR} &= r\beta d^{EPR} + \ln \left[\frac{\Phi \left(\frac{d^{EPR}}{\theta} + \frac{r\beta\theta}{2} \right)}{\Phi \left(-\frac{d^{EPR}}{\theta} + \frac{r\beta\theta}{2} \right)} \right]\end{aligned}\quad (42)$$

The second term (in curly brackets) in (41) has the sign of

$$-\ln \left[2\Phi \left(\frac{d^{EPR}}{\theta} + \frac{r\beta\theta}{2} \right) \right] - \frac{r\beta\theta}{2} \frac{d^{EPR}}{\theta} 2\Phi \left(-\frac{d^{EPR}}{\theta} \right) + \frac{r\beta\theta}{2} \frac{d^{EAR}}{\theta} + \frac{r\beta\theta}{2} 2\phi \left(\frac{d^{EPR}}{\theta} \right). \quad (43)$$

Define $t \equiv \frac{r\beta\theta}{2}$, $y^{EPR} \equiv \frac{d^{EPR}}{\theta}$, $y^{EAR} \equiv \frac{d^{EAR}}{\theta}$. Then (43) can be rewritten as

$$-\ln [2\Phi (y^{EPR} + t)] - ty^{EPR} 2\Phi (-y^{EPR}) + ty^{EAR} + 2t\phi (y^{EPR}) .$$

Using this notation to rewrite (42) we have

$$ty^{EAR} = ty^{EPR} + \frac{1}{2} \ln \left[\frac{\Phi (y^{EPR} + t)}{\Phi (-y^{EPR} + t)} \right] .$$

Using this equation to substitute into the one above we obtain

$$-\ln [2\Phi (y^{EPR} + t)] + \frac{1}{2} \ln \left[\frac{\Phi (y^{EPR} + t)}{\Phi (-y^{EPR} + t)} \right] + ty^{EPR} [2\Phi (y^{EPR}) - 1] + 2t\phi (y^{EPR}) .$$

This expression has the same sign as

$$f(y^{EPR}, t) \equiv -\ln [4\Phi (y^{EPR} + t) \Phi (-y^{EPR} + t)] + 2ty^{EPR} [2\Phi (y^{EPR}) - 1] + 4t\phi (y^{EPR})$$

Now of course y^{EPR} depends on t (and λ , which does not appear explicitly here) through equation (15). However, we can ignore this dependence because we can show that $\forall t > 0, \forall y^{EPR} \geq 0$, the function $f(y^{EPR}, t)$ defined above is strictly positive.

Lemma 2 $\forall t > 0, \forall y^{EPR} \geq 0, f(y^{EPR}, t) > 0$.

Proof of Lemma 2. (i) If $y^{EPR} = 0$, then

$$\begin{aligned} f(0, t) &= -\ln[4(\Phi(t))^2] + 4t\phi(0) \\ &= 2\{-\ln[2\Phi(t)] + 2t\phi(0)\} \end{aligned}$$

and so

$$f(0, 0) = 0$$

as well as

$$\begin{aligned} \frac{\partial f(0, t)}{\partial t} &= 2 \left[-\frac{\phi(t)}{\Phi(t)} + 2\phi(0) \right] \\ &= 2 \left[-\frac{\phi(t)}{\Phi(t)} + \frac{\phi(0)}{\Phi(0)} \right] \\ &> 0 \end{aligned}$$

since $\frac{\phi(t)}{\Phi(t)}$ is decreasing in t , $\forall t$. Therefore, $f(0, t) > 0, \forall t > 0$.

(ii) The partial derivative of $f(y^{EPR}, t)$ with respect to y^{EPR} is given by

$$\begin{aligned} \frac{\partial f(y^{EPR}, t)}{\partial y^{EPR}} &= -\frac{\phi(y^{EPR} + t) \Phi(-y^{EPR} + t) - \Phi(y^{EPR} + t) \phi(-y^{EPR} + t)}{\Phi(y^{EPR} + t) \Phi(-y^{EPR} + t)} \\ &\quad + 2t [2\Phi(y^{EPR}) - 1] + 4ty^{EPR} \phi(y^{EPR}) + 4t\phi'(y^{EPR}) \\ &= \frac{\phi(-y^{EPR} + t)}{\Phi(-y^{EPR} + t)} - \frac{\phi(y^{EPR} + t)}{\Phi(y^{EPR} + t)} + 2t [2\Phi(y^{EPR}) - 1] \end{aligned}$$

since $\phi'(y^{EPR}) = -y^{EPR}\phi(y^{EPR})$. Thus, $\forall t > 0, \forall y^{EPR} \geq 0$

$$\frac{\partial f(y^{EPR}, t)}{\partial y^{EPR}} \geq 0.$$

(i) and (ii) together imply that $\forall t > 0, \forall y^{EPR} \geq 0, f(y^{EPR}, t) > 0$. ■

It follows from the lemma that, for the agent's optimal value of y^{EPR} and the corresponding y^{EAR} , given any $t > 0, f(y^{EPR}, t) > 0$. Since $t \equiv \frac{r\beta\theta}{2}$ and $\theta \equiv \sigma(1 - rho)^{\frac{1}{2}}$, $t > 0$ if $\beta > 0$ and $\rho < 1$. We have thus proved that the costs of compensating the agent are lower under ex post than under ex ante randomization, $\forall \beta > 0, \forall \rho < 1$. For $\rho = 1$, it can be proved directly, using Propositions 1 and 2, that the second term (in curly brackets) in (41) equals 0, so in this case the costs of compensating the agent are the same under the two types of randomization.

Overall, therefore, for any $\beta > 0$, ex post randomization generates higher expected benefits for the principal than ex ante randomization if $\delta \geq \lambda$, and lower costs of compensating the agent (strictly lower costs if $\rho < 1$). This proves the proposition. ■

Proof of Proposition 4. Given the analysis in the text immediately following the statement of Proposition 4, it remains only to prove that $\forall \beta > 0, \Pi^{SD}\left(\frac{\beta}{2}\right) - \Pi^{EPR}(\beta) \geq 0$, with strict inequality when $\rho < 1$. From equation (20),

$$\Pi^{SD}\left(\frac{\beta}{2}\right) - \Pi^{EPR}(\beta) = \frac{1}{4}R\beta^2(1 - \rho) + \frac{1}{r} \ln \left[2\Phi\left(\frac{r\beta\sigma\sqrt{1 - \rho}}{\sqrt{2}}\right) \right] - \frac{\beta\sigma\sqrt{1 - \rho}}{\sqrt{\pi}}.$$

Now use the definition $\theta = \sqrt{2}\sigma\sqrt{1 - \rho}$, so $\sigma\sqrt{1 - \rho} = \frac{\theta}{\sqrt{2}}$. With this, we have

$$\Pi^{SD}\left(\frac{\beta}{2}\right) - \Pi^{EPR}(\beta) = \frac{1}{r} \left\{ \frac{1}{2} \left(\frac{r\beta\theta}{2} \right)^2 - \frac{r\beta\theta}{2} \sqrt{\frac{2}{\pi}} + \ln \left[2\Phi\left(\frac{r\beta\theta}{2}\right) \right] \right\}$$

We can write the factor in the curly brackets in terms of $t \equiv \frac{r\beta\theta}{2}$ that is

$$g(t) \equiv \frac{t^2}{2} - \sqrt{\frac{2}{\pi}}t + \ln [2\Phi(t)].$$

Analyzing this function we have

$$\begin{aligned} g(0) &= 0 \\ g'(t) &= t - \sqrt{\frac{2}{\pi}} + \frac{\phi(t)}{\Phi(t)} = -\sqrt{\frac{2}{\pi}} + \frac{t\Phi(t) + \phi(t)}{\Phi(t)} \\ g'(0) &= 0 \\ g''(t) &= \frac{[\Phi(t) + t\phi(t) + \phi'(t)]\Phi(t) - [t\Phi(t) + \phi(t)]\phi(t)}{[\Phi(t)]^2} \\ &= \frac{[\Phi(t)]^2 - t\phi(t)\Phi(t) - [\phi(t)]^2}{[\Phi(t)]^2} \\ g''(0) &= \frac{1}{4} - \frac{1}{2\pi} > 0 \end{aligned}$$

and finally the derivative of the numerator of $g''(t)$ is

$$\begin{aligned}\frac{\partial}{\partial t} \left\{ [\Phi(t)]^2 - t\phi(t)\Phi(t) - [\phi(t)]^2 \right\} &= 2\Phi\phi - \phi\Phi - t\phi'\Phi - t\phi^2 - 2\phi\phi' \\ &= \phi\Phi + t^2\phi\Phi - t\phi^2 + 2t\phi^2 \\ &> 0\end{aligned}$$

for $t > 0$. Therefore,

$$\begin{aligned}\forall t &> 0, \quad g''(t) > 0 \\ \forall t &> 0, \quad g'(t) > 0 \\ \forall t &> 0, \quad g(t) > 0\end{aligned}$$

Hence, since $\beta > 0$ and $\rho < 1$ imply that $t > 0$, we have shown that $\forall \beta > 0$ and $\rho < 1$, $\Pi^{SD}\left(\frac{\beta}{2}\right) - \Pi^{EPR}(\beta) > 0$. If $\rho = 1$, then $\theta = 0$ and $t = 0$, so $\Pi^{SD}\left(\frac{\beta}{2}\right) - \Pi^{EPR}(\beta) = 0$. ■

Proof of Proposition 5. We begin by showing that SD dominates EAR under the conditions of the proposition. When, for a given β , EAR induces a corner solution, \bar{e}^{EAR} is given by the FOC

$$\frac{\beta - \bar{e}^{EAR}}{\bar{e}^{EAR}} = \exp \{ r\beta \bar{e}^{EAR} \}. \quad (44)$$

Since the RHS of (44) is ≥ 1 for $\bar{e}^{EAR} \geq 0$,

$$\bar{e}^{EAR} \leq \beta/2. \quad (45)$$

When EAR induces A to choose the corner solution $(\bar{e}^{EAR}, 0)$,

$$\Pi^{EAR}(\beta, \bar{e}^{EAR}) = \frac{\bar{e}^{EAR}}{\delta} - \frac{1}{2}(\bar{e}^{EAR})^2 - \frac{1}{2}r\sigma^2(\beta)^2 - \frac{1}{2r} \ln \left(\frac{(1+Z)^2}{4Z} \right),$$

where $Z = \frac{\beta - \bar{e}^{EAR}}{\bar{e}^{EAR}} \geq 1$. To induce the same effort pair $(\bar{e}^{EAR}, 0)$ under SD requires $\beta^{SD} = \bar{e}^{EAR}$ and thus

$$\Pi^{SD}(\beta^{SD}, \bar{e}^{EAR}) = \frac{\bar{e}^{EAR}}{\delta} - \frac{1}{2}(\bar{e}^{EAR})^2 - \frac{1}{2}r\sigma^2(1+\rho)^2(\bar{e}^{EAR})^2.$$

Therefore

$$\begin{aligned}\Pi^{SD}(\beta^{SD}, \bar{e}^{EAR}) - \Pi^{EAR}(\beta, \bar{e}^{EAR}) &\geq \frac{r\sigma^2}{2}(\beta^2 - 2(1+\rho)\bar{e}^{EAR})^2 \\ &\geq \frac{r\sigma^2}{2}(4 - 2(1+\rho))(\bar{e}^{EAR})^2 \\ &\geq 0,\end{aligned}$$

where the first inequality holds since $Z \geq 1$ and the second since $\beta \geq 2\bar{e}^{EAR}$ by (45).

We now compare SD and EPR. When, for a given β , EPR induces a corner solution, \bar{e}^{EPR} is given by the FOC

$$\frac{\beta - \bar{e}^{EPR}}{\bar{e}^{EPR}} = \exp \{ r\beta \bar{e}^{EPR} \} \frac{\Phi \left(\frac{\bar{e}^{EPR} + r\beta\sigma^2(1-\rho)}{\theta} \right)}{\Phi \left(\frac{-\bar{e}^{EPR} + r\beta\sigma^2(1-\rho)}{\theta} \right)}. \quad (46)$$

Since the RHS of (46) is ≥ 1 for $\bar{e}^{EPR} \geq 0$, $\bar{e}^{EPR} \leq \beta/2$. When EPR induces A to choose the corner solution $(\bar{e}^{EPR}, 0)$,

$$\begin{aligned}\Pi^{EPR}(\beta, \bar{e}^{EPR}) &= \frac{\bar{e}^{EPR}}{\delta} - \frac{1}{2}(\bar{e}^{EPR})^2 - \frac{1}{2}r\beta^2\sigma^2 - \frac{1}{r}\ln\left(\Phi\left(\frac{\bar{e}^{EPR} + r\beta\sigma^2(1-\rho)}{\theta}\right)\right) \\ &\quad + \exp\{-r\beta\bar{e}^{EPR}\}\Phi\left(\frac{-\bar{e}^{EPR} + r\beta\sigma^2(1-\rho)}{\theta}\right).\end{aligned}$$

To induce the same effort pair $(\bar{e}^{EPR}, 0)$ under SD requires $\beta^{SD} = \bar{e}^{EPR}$ and thus

$$\Pi^{SD}(\beta^{SD}, \bar{e}^{EPR}) = \frac{\bar{e}^{EPR}}{\delta} - \frac{1}{2}(\bar{e}^{EPR})^2 - (1+\rho)r\sigma^2(\bar{e}^{EPR})^2.$$

Therefore, denoting $y^{EPR} \equiv \bar{e}^{EPR}/\theta$ and $t \equiv r\beta\theta/2$, we can write

$$\begin{aligned}\Pi^{SD}(\beta^{SD}, \bar{e}^{EPR}) - \Pi^{EPR}(\beta, \bar{e}^{EPR}) &= \frac{r\sigma^2}{4}\left(2(\beta)^2 - 4(1+\rho)(\bar{e}^{EPR})^2\right) - \frac{1}{r}(h(y^{EPR}, t)) \\ &\geq \frac{r\sigma^2}{4}\left(2(\beta)^2 - 4(1+\rho)(\bar{e}^{EPR})^2\right) - \frac{1}{r}(h(0, t)) \\ &\stackrel{sgn}{=} \frac{r^2\sigma^2}{4}\left(2(\beta)^2 - 4(1+\rho)(\bar{e}^{EPR})^2\right) - (h(0, t)) \\ &\geq \frac{r^2\sigma^2}{4}\left(2(\beta)^2 - (1+\rho)(\beta)^2\right) - (h(0, t)) \\ &= \frac{r^2\sigma^2(\beta)^2(1-\rho)}{4} - (h(0, t)) \\ &= \frac{(t)^2}{2} + \ln(2\Phi(t)) - 2t\phi(0) \geq 0, \forall t > 0.\end{aligned}$$

where we used the fact that $\beta \geq 2\bar{e}^{EPR}$ and the fact that $h(y, t)$ is decreasing in y .¹⁶ Note also that the last line was established in the proof of Proposition 4. ■

Proof of Proposition 6. There are two cases to consider: (i) EAR and EPR induce interior efforts or (ii) EAR and EPR induce corner solutions. Proposition 5 deals with the latter case, so here we treat the former.

For ex ante randomization, the principal's profit is

$$\Pi^{EAR}(\beta) = \underline{e} + \frac{1}{\delta}\bar{e} - \frac{\ln\left[\frac{(\lambda+1)^2}{4\lambda}\right]}{2r} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2\beta^2.$$

Note that

$$\Pi^{EAR}(\beta) \leq \frac{\beta}{\delta(\lambda+1)} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2\beta^2$$

since

$$\underline{e} + \frac{1}{\delta}\bar{e} = \frac{1}{\delta}(\delta\underline{e} + \bar{e}) \leq \frac{1}{\delta}(\lambda\underline{e} + \bar{e}) = \frac{\beta}{\delta(\lambda+1)}$$

¹⁶This is the result that the risk costs imposed on the agent by ex post randomization are increasing in the gap $\bar{e} - \underline{e}$, reflecting the fact that the variance of $w = \min\{\alpha + \beta x_1, \alpha + \beta x_2\}$ is increasing in $\bar{e} - \underline{e}$. See the footnote at the end of the discussion after Proposition 2.

given the assumption that optimal efforts are interior. By contrast, the profits for the SD contract with incentive coefficient β are

$$\Pi^{SD}(\beta) = \frac{\beta}{\delta} - \frac{(\beta)^2}{2} - r\sigma^2(1+\rho)\beta^2.$$

Setting $\beta^{SD} = \frac{\beta}{\lambda+1}$, we have

$$\Pi^{SD}\left(\frac{\beta}{\lambda+1}\right) = \frac{\beta}{\delta(\lambda+1)} - \frac{\beta^2}{2(\lambda+1)^2} - r\sigma^2(1+\rho)\frac{\beta^2}{(\lambda+1)^2}$$

and

$$\begin{aligned} \Pi^{SD}\left(\frac{\beta}{\lambda+1}\right) - \Pi^{EAR}(\beta) &\geq r\sigma^2\beta^2 \left[\frac{1}{2} - \frac{1+\rho}{(\lambda+1)^2} \right] \\ &\geq 0. \end{aligned}$$

Now consider ex post randomization. Given β , $\bar{e} + \lambda\underline{e} = \beta/(\lambda+1)$. Define $d \equiv \bar{e} - \underline{e}$. Then from Proposition 2, the principal's profits are given by

$$\begin{aligned} \Pi^{EPR}(\beta) &= \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 - \frac{1}{2}r\beta^2\sigma^2 \\ &\quad - \frac{1}{r} \ln \left[\exp\{-r\beta\bar{e}\} \Phi\left(\frac{-d + r\beta\sigma^2(1-\rho)}{\theta}\right) + \exp\{-r\beta\underline{e}\} \Phi\left(\frac{d + r\beta\sigma^2(1-\rho)}{\theta}\right) \right] \\ &\quad - \beta d \Phi\left(\frac{-d + r\beta\sigma^2(1-\rho)}{\theta}\right) + \beta\theta\phi\left(\frac{d}{\theta}\right). \end{aligned}$$

Consider the final two lines of the above. They are equal to

$$-\frac{1}{r} [\ln [\Phi(y+t) + \exp\{-2ty\} \Phi(-y+t)] - 2ty\Phi(-y)] + 2t\phi(-y) \equiv \frac{1}{r} \{h(y, t)\},$$

where $y \equiv \frac{d}{\theta}$ and $t \equiv \frac{r\beta\theta}{2}$. We have previously shown that $h(y, t)$ is decreasing in y for $y \geq 0, t \geq 0$.¹⁷ Therefore, replacing $h(y, t)$ by $h(0, t)$ yields

$$\begin{aligned} \Pi^{EPR}(\beta) &\leq \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 - \frac{1}{2}r\beta^2\sigma^2 + \frac{1}{r} [-\ln(2\Phi(t)) + 2t\phi(0)] \\ &\leq \frac{\bar{e} + \lambda\underline{e}}{\delta} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 - \frac{1}{2}r\beta^2\sigma^2 + \frac{1}{r} [-\ln(2\Phi(t)) + 2t\phi(0)] \\ &= \frac{\beta}{\delta(1+\lambda)} - \frac{1}{2}\frac{\beta^2}{(1+\lambda)^2} - \frac{1}{2}r\beta^2\sigma^2 + \frac{1}{r} [-\ln(2\Phi(t)) + 2t\phi(0)], \end{aligned}$$

where the second line uses the fact that $\delta \leq \lambda$. Setting $\beta^{SD} = \beta/(1+\lambda)$ and using the definition

¹⁷This is, again, the result that the risk costs imposed on the agent by ex post randomization are increasing in the gap $\bar{e} - \underline{e}$, reflecting the fact that the variance of $w = \min\{\alpha + \beta x_1, \alpha + \beta x_2\}$ is increasing in $\bar{e} - \underline{e}$.

of $t \equiv r\beta^{EPR}\theta$, we have

$$\begin{aligned}
\Pi^{SD}\left(\frac{\beta}{1+\lambda}\right) - \Pi^{EPR}(\beta) &= r\beta^2\sigma^2\left(\frac{1}{2} - \frac{(1+\rho)}{(1+\lambda)^2}\right) - \frac{1}{r}\left[-\ln\left(2\Phi\left(\frac{r\beta\theta}{2}\right)\right) + r\beta\theta\phi(0)\right] \\
&\stackrel{sgn}{=} (r\beta\sigma)^2\left(\frac{1}{2} - \frac{(1+\rho)}{(1+\lambda)^2}\right) + \ln(2\Phi(t)) - 2t\phi(0) \\
&\geq (r\beta\sigma)^2\left(\frac{1}{2} - \frac{(1+\rho)}{4}\right) + \ln(2\Phi(t)) - 2t\phi(0) \\
&= \frac{(r\beta\sigma)^2}{4}(2 - (1+\rho)) + \ln(2\Phi(t)) - 2t\phi(0) \\
&= \frac{t^2}{2} + \ln(2\Phi(t)) - 2t\phi(0) \geq 0, \forall t \geq 0,
\end{aligned}$$

where the third line uses the fact that $\lambda \geq 1$ and the final line was established in the proof of Proposition 4. ■

Proof of Corollary 2. Under ex ante randomization, both types of agent will choose a corner solution for efforts whenever

$$\beta^2 \leq \frac{(1+\lambda)\ln\lambda}{r}.$$

First we will show that the condition given in the statement of the corollary implies that the optimal β , derived ignoring the non-negativity constraint on \underline{e} , is such that it induces a corner solution. Then we will confirm that under this condition the optimal ex ante randomization scheme, taking the non-negativity constraint on \underline{e} into account, does indeed induce a corner solution for efforts.

If we ignore the non-negativity constraint on efforts and assume that the agent chooses (\bar{e}, \underline{e}) to satisfy both of the first-order conditions with equality, then the derivative with respect to β of the principal's expected profit under ex ante randomization has the same sign as

$$\frac{\delta+1}{\delta} + \left(\frac{\delta-\lambda}{\delta}\right) \frac{(1+\lambda)\ln\lambda}{r\beta^2} - \beta \left(1 + r\sigma^2(1+\lambda)^2\right). \quad (47)$$

This expression is strictly decreasing in β , and there is a unique value of β , $\hat{\beta}$, at which it equals 0. We will have

$$(\hat{\beta})^2 \leq \frac{(1+\lambda)\ln\lambda}{r}$$

if and only if (47) evaluated at $\beta = ((1+\lambda)\ln\lambda)/r$ is non-positive, i.e.

$$\begin{aligned}
\frac{\delta+1}{\delta} + \frac{\delta-\lambda}{\delta} - \left(1 + r\sigma^2(1+\lambda)^2\right) \sqrt{\frac{(1+\lambda)\ln\lambda}{r}} &\leq 0 \\
\Leftrightarrow 2 - \left(\frac{\lambda-1}{\delta}\right) &\leq \left(1 + r\sigma^2(1+\lambda)^2\right) \sqrt{\frac{(1+\lambda)\ln\lambda}{r}} \\
\Leftrightarrow 2 - J &\leq \frac{\lambda-1}{\delta}. \quad (48)
\end{aligned}$$

If $J > 2$, the condition (48) is satisfied for all $\delta > 0$. Since (47) is strictly decreasing in β , it follows that whenever (47) evaluated at $\beta = ((1+\lambda)\ln\lambda)/r$ is non-positive, the principal's profit under ex ante randomization would be strictly lowered by raising β from $((1+\lambda)\ln\lambda)/r$, the largest value which induces a corner solution for efforts, to a higher value and so inducing an interior solution for

efforts. Hence, when $J > 2$, the optimal ex ante randomization scheme induces a corner solution for efforts for both types of agent. ■

Proof of Lemma 1. For $\rho < 1$, we derived

$$\bar{e} + \lambda \underline{e} = \frac{\beta}{\lambda + 1}$$

and, with $d \equiv \bar{e} - \underline{e}$,

$$\lambda = \exp[r\beta d] \frac{\Phi\left(\frac{d+r\beta\sigma^2(1-\rho)}{\sigma\sqrt{2(1-\rho)}}\right)}{\Phi\left(\frac{-d+r\beta\sigma^2(1-\rho)}{\sigma\sqrt{2(1-\rho)}}\right)} \quad (49)$$

First, we claim that

$$\lim_{\rho \rightarrow 1} d(\rho) = 0.$$

To see this, note that as $\rho \rightarrow 1$, equation (49) becomes

$$\lambda = \exp[r\beta d(1)] \lim_{\rho \rightarrow 1} \frac{\Phi\left(\frac{d(\rho)}{\sigma\sqrt{2(1-\rho)}}\right)}{\Phi\left(\frac{-d(\rho)}{\sigma\sqrt{2(1-\rho)}}\right)}. \quad (50)$$

If $d(1)$ were to equal a strictly positive number, then

$$\lim_{\rho \rightarrow 1} \frac{\Phi\left(\frac{d(\rho)}{\sigma\sqrt{2(1-\rho)}}\right)}{\Phi\left(\frac{-d(\rho)}{\sigma\sqrt{2(1-\rho)}}\right)} \rightarrow \infty,$$

which would mean that equation (50) was not satisfied. Only if $\lim_{\rho \rightarrow 1} d(\rho) = 0$ can equation (50) be satisfied. Thus in the limit as $\rho \rightarrow 1$, $d(\rho) \rightarrow 0$, and $\frac{d(\rho)}{\theta}$ approaches the solution, k^* , to

$$\lambda = \frac{\Phi(k)}{\Phi(-k)}.$$

Since $\bar{e} + \lambda \underline{e}$ is independent of ρ , we conclude that

$$\lim_{\rho \rightarrow 1} \bar{e} = \lim_{\rho \rightarrow 1} \underline{e} = \frac{\beta}{(\lambda + 1)^2}.$$

Thus the optimal behavior of the agent is continuous at $\rho = 1$. The expected profit of the principal is given by

$$\begin{aligned} \Pi^{EPR} &= \underline{e}^{EPR} + \frac{1}{\delta} \bar{e}^{EPR} - \frac{\beta^2}{2(\lambda + 1)^2} - \frac{1}{2} r \sigma^2 \beta^2 \\ &\quad - \frac{1}{r} \ln \left[\frac{1 + \lambda}{\lambda} \Phi\left(\frac{d + r\beta\sigma^2(1-\rho)}{\theta}\right) \right] - \beta d \Phi\left(-\frac{d}{\theta}\right) + \beta \theta \phi\left(\frac{d}{\theta}\right) \end{aligned}$$

which using equation (49) we can rewrite as

$$\begin{aligned}
\Pi^{EPR} &= \underline{e}^{EPR} + \frac{1}{\delta} \bar{e}^{EPR} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2} r \sigma^2 \beta^2 \\
&\quad - \frac{1}{r} \ln \left[\Phi \left(\frac{d + r \beta \sigma^2 (1 - \rho)}{\theta} \right) + \exp[-r \beta d] \Phi \left(\frac{-d + r \beta \sigma^2 (1 - \rho)}{\theta} \right) \right] \\
&\quad - \beta d \Phi \left(-\frac{d}{\theta} \right) + \beta \theta \phi \left(\frac{d}{\theta} \right).
\end{aligned}$$

Taking the limit as $\rho \rightarrow 1$, we obtain

$$\lim_{\rho \rightarrow 1} \Pi^{EPR}(\beta, \rho) = \frac{\delta+1}{\delta} \frac{\beta}{(\lambda+1)^2} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{r \sigma^2 \beta^2}{2} - \frac{1}{r} \ln [\Phi(k^*) + \Phi(-k^*)]$$

since

$$\begin{aligned}
\lim_{\rho \rightarrow 1} \beta d(\rho) \Phi \left(-\frac{d}{\theta} \right) &= 0 \\
\lim_{\rho \rightarrow 1} \beta \theta \phi \left(\frac{d}{\theta} \right) &= 0
\end{aligned}$$

because $d(\rho) \rightarrow 0$ and $\theta \rightarrow 0$. Finally, observe that

$$\Phi(k^*) + \Phi(-k^*) = 1,$$

so

$$\lim_{\rho \rightarrow 1} \Pi^{EPR}(\beta, \rho) = \frac{\delta+1}{\delta} \frac{\beta}{(\lambda+1)^2} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{r \sigma^2 \beta^2}{2}.$$

This, coupled with the direct derivation in the text of equation (53) for $\rho = 1$ shows that the principal's profit under ex post randomization is continuous at $\rho = 1$. ■

Proof of Proposition 8. The claim in Part 1 follows straightforwardly from comparing the profit expressions for SD and EPR as $\rho \rightarrow 1$ and $R \rightarrow 0$.

To prove Part 2, first note that for $\rho = 1$ and $R \in (0, 1/2]$, $\rho R \leq 1/2$, so the optimal deterministic scheme is as given in the second panel of Figure 1 (OA is dominated by SD). Now show, by comparing the profit expressions for OT and EPR that there exists a critical $\delta^{OT/EPR}$ above (below) which EPR dominates (is dominated by) OT; by comparing the profit expressions for SD and EPR that there exists a critical $\delta^{SD/EPR}$ above (below) which EPR dominates (is dominated by) SD; by comparing the profit expressions for AD and EPR that there exists a critical $\delta^{AD/EPR}$ above (below) which EPR dominates (is dominated by) AD; and, by comparing the profit expressions for SD and AD that there exists a critical $\delta^{SD/AD}$ above (below) which AD dominates SD. For $R \leq 1/2$, we can show that $\delta^{SD/EPR} < \delta^{SD/AD}$ for all $\lambda > 1$. This inequality, coupled with the existence of the critical $\delta^{AD/EPR}$, implies that for all $\lambda > 1$ and $R \leq 1/2$, AD is dominated by one or both of SD and EPR, so AD can never be the most profitable scheme when $R \leq 1/2$. Comparison of the profit expressions for SD and OT shows that there is a critical $\hat{\lambda}(R)$ above (below) which SD dominates (is dominated by) OT and that $\hat{\lambda}(R)$ is increasing in R , going to 1 as $R \rightarrow 0$ and to 1.68 as $R \rightarrow 1/2$.

To prove Part 3, first note that for $\rho = 1$ and $R \rightarrow \infty$, $\rho R > 1/2$, so the optimal deterministic scheme is as given in the third panel of Figure 1 (SD is dominated by OA). Comparison of the profit expressions for OT and OA shows that for $\lambda \leq 1 + \sqrt{2}$, $\Pi^{OA} \geq \Pi^{OT}$. Also show, by comparing

the profit expressions for OA and EPR, that there exists a critical $\delta^{OA/EPR}$ above (below) which EPR dominates (is dominated by) OA and, by comparing the profit expressions for OA and AD, that there exists a critical $\delta^{OA/AD}$ above (below) which AD dominates (is dominated by) OA. For $\lambda \leq 1 + \sqrt{2}$, $\delta^{OT/EPR} < \delta^{OT/AD}$, and this inequality, coupled with the existence of the critical $\delta^{AD/EPR}$, implies that for $\lambda \leq 1 + \sqrt{2}$, AD is dominated by one or both of OT and EPR, so AD cannot be the most profitable scheme. On the other hand, for $\lambda > 1 + \sqrt{2}$, $\delta^{OA/AD} < \delta^{AD/EPR}$, so for $\delta > \delta^{AD/EPR}$, EPR is the optimal scheme, for $\delta \in (\delta'', \delta')$ where $\delta''^{OA/AD}$, AD is the optimal scheme, and for $\delta < \delta''$, OA is the optimal scheme. ■

Proof of Proposition 12. Assume $\rho < 1$, and consider the limiting case where $r \rightarrow 0$ and $\sigma^2 \rightarrow \infty$ in such a way that $r\sigma^2 \rightarrow R \geq 0$. The right-hand side of the agents' first-order condition in Proposition 11 converges to 1, so $\bar{e}^* \rightarrow \beta/2$. Applying L'Hôpital's Rule to the principal's profit expression (33) in Proposition 11 yields

$$\Pi^{\max}(\beta) \rightarrow \frac{\beta}{2\delta} - \frac{\beta^2}{8} - \frac{1}{2}R\beta^2 \left(1 - \frac{1-\rho}{\pi}\right).$$

Since

$$\Pi^{SD}\left(\frac{\beta}{2}\right) = \frac{\beta}{2\delta} - \frac{\beta^2}{8} - \frac{1}{4}R\beta^2(1+\rho),$$

and $\Pi^{\max}(\beta) \leq \Pi^{SD}\left(\frac{\beta}{2}\right)$, it follows that SD dominates the max contract in this limiting case. ■

Proof of Proposition 13. Part 2(iv) of Proposition 11 shows that \bar{e}^* is continuous at $\rho = 1$ and that for $\rho = 1$, $\bar{e}^* = \beta$. The principal's profit expression (33) in Proposition 11 is also continuous at $\rho = 1$, and taking the limit in (33) as $\rho \rightarrow 1$ yields

$$\Pi^{\max}(\beta) = \frac{\beta}{\delta} - \frac{1}{2}\beta^2 - \frac{1}{2}R\beta^2.$$

Therefore maximized profits under the max contract are

$$\Pi^{\max} = \frac{1}{2\delta^2(1+R)}. \quad (51)$$

Comparing equation (51) with equations (4), (5), and (6) giving the maximized profits for the SD, OT, and OA schemes, respectively, shows that for $\rho = 1$, the max contract dominates each of these other three contracts, for all $\lambda > 1$, for all $\delta \geq 1$, and for all $R \geq 0$. This establishes part 1.

To establish part 2, recall from equations (8) and (26) that when $\rho = 1$, maximized profits under AD and EPR are

$$\Pi^{AD} = \frac{(\lambda^2 + \lambda + \delta + 1)^2}{8\delta^2(1+\lambda)^2(\lambda^2 + (1+\lambda)^2 R)}, \quad (52)$$

$$\Pi^{EPR} = \frac{(\delta + 1)^2}{2\delta^2(1+\lambda)^2(1 + (1+\lambda)^2 R)}. \quad (53)$$

Using these expressions, along with (51), we can show that for $\delta \leq \delta^{\max/EPR}$, $\Pi^{EPR} \leq \Pi^{\max}$, where

$$\delta^{\max/EPR} = (1+\lambda) \left(\frac{1 + (1+\lambda)^2 R}{1+R} \right) - 1.$$

Also, for $\delta \leq \delta^{\max/AD}$, $\Pi^{AD} \leq \Pi^{\max}$, where

$$\delta^{\max/AD} = 2(1 + \lambda) \left(\frac{\lambda^2 + (1 + \lambda)^2 R}{1 + R} \right)^{\frac{1}{2}} - (\lambda^2 + \lambda + 1).$$

Furthermore, we can show that $\delta^{\max/EPR} < \delta^{\max/AD}$ for all $\lambda > 1$ and $R \geq 0$. This inequality, coupled with the fact that there exists a critical $\delta^{AD/EPR}$ such that

$$\delta \leq \delta^{AD/EPR} \Leftrightarrow \Pi^{EPR} \leq \Pi^{AD},$$

implies that, for all $\lambda > 1$ and $R \geq 0$, if $\delta < \delta^{\max/EPR}$, then the max contract dominates AD and EPR, and if $\delta > \delta^{\max/EPR}$, then EPR dominates AD and the max contract. ■