Gaming and Strategic Ambiguity in Incentive Provision[*]

Florian Ederer[†]   Richard Holden[‡]   Margaret Meyer[§]
MIT   MIT and NBER   Oxford and CEPR

April 7, 2009

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)

[†]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

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 was marked by significant ambiguity about the criteria that were to be used, 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)

[2] See 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, and only the agent knows which type he is. One type has a lower cost of effort on task 1, and the other 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. However, we show that the agent's private information about his preferences makes such an outcome 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.

Ex ante randomization pushes the agent toward balanced efforts on the tasks as a means of partially 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 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.

In addition to the simple forms of random contracts just described, we also analyze more general forms of ex ante and ex post randomization, in which the two possible compensation schemes that might ultimately be used both assign positive weights to both performance measures, but they differ with respect to which indicator is more highly rewarded. We show how the principal can use the relative weight on the two performance measures to adjust the intensity of the incentives provided to the agent to choose balanced efforts.

The benefits of randomized incentive schemes in deterring gaming do, nevertheless, come at

---

[3]For 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.

a cost in terms of the risk they impose on the agent. Specifically, we show that a deterministic contract can induce any given level of *aggregate* effort on the two tasks while imposing lower risk costs than any randomized scheme. In general, therefore, the principal faces a trade-off between the stronger incentives for balanced efforts that arise under randomized schemes and the better insurance that is provided by deterministic schemes.

Our key contribution is to identify settings in which optimally designed randomized contracts dominate all deterministic incentive schemes. We identify three such environments. Each of these environments has the feature that optimally designed randomized contracts induce both types of 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 the limiting case where the agents' risk aversion becomes infinitely large and the variance of the shocks to outputs becomes arbitrarily small. The final setting is that where the shocks affecting measured performance on the tasks become perfectly correlated. In all three of these environments, we show that there is a critical degree of complementarity of tasks for the principal above which the optimal incentive scheme is a randomized one.

As the above discussion foreshadowed, our model can best be thought of in the light of two path-breaking 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 limitations on the principal's knowledge of the contracting environment.[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 Malcomson 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

---

[4]A 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).

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 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 risk tolerance varies with the level of effort he exerts. 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 his action. A *sufficient* condition for this is that the agent's utility function is additively or multiplicatively separable in action and reward. 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 that therefore, 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 contacts *cannot* dominate the best deterministic contract. Thus, the attractiveness of random schemes arises because of the agent's 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

---

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

advance. Under ex ante randomization, the outcome is equivalent to the equilibrium outcome of a game in which the principal chooses the randomizing probability at the same time as the agent chooses efforts. Ex post randomization allows the principal to retain discretion over which performance measure to reward until after observing outputs. 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 with limited commitment.*

The remainder of the paper is organized as follows. Section 2 outlines the model. In section 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 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. 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 measured performance, 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 substitutes[6]: $\frac{\partial c_i/\partial e_1}{\partial c_i/\partial e_2}$ 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

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

---

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

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

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

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 performance measures), and $R \equiv r\sigma^2$ (which represents the importance of risk aversion).[7]

We then introduce the two classes of randomized contracts on which we focus. We begin with *ex ante randomization*. In the simplest form of ex ante randomization (EAR), the contract specifies that with probability $\frac{1}{2}$, $\beta_1 = \beta$ and $\beta_2 = 0$ and with probability $\frac{1}{2}$, $\beta_1 = 0$ and $\beta_2 = \beta$. Under this scheme, the principal commits to employ a randomizing device to determine on which of the two outputs the agent's pay will be based. Under a more general form of ex ante randomization (gEAR), the agent is compensated according to $w = \alpha + \beta x_1 + k\beta x_2$ with probability $\frac{1}{2}$ and according to $w = \alpha + \beta x_2 + k\beta x_1$ with probability $\frac{1}{2}$, where $k \in (-1, 1)$. The second class of randomized contract is termed *ex post randomization*. In its simplest form (EPR), the principal, after observing the outputs x_1 and x_2, chooses whether to pay the agent $\alpha + \beta x_1$ or $\alpha + \beta x_2$. In its more general form (gEPR), the principal chooses between paying $w = \alpha + \beta x_1 + k\beta x_2$ or $w = \alpha + \beta x_2 + k\beta x_1$, where again $k \in (-1, 1)$. Under both classes of randomized incentive schemes, the agent is ex ante uncertain about what weights the two performance indicators will be given in determining his pay, but only under ex post randomization can the agent's efforts influence which set of weights is ultimately used.

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 on only one task and uses performance on the other task to provide insurance for the agent, by exploiting the correlation between the shocks to the performance measures. This type of contract, which we denote by OT (for "one task") is of the form

$$w = \alpha + \beta x_1 - \rho\beta x_2.$$

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

---

[7]For 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.

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

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

With β chosen optimally, the resulting maximized profit is

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

Under the OT contract, $\beta_1 = \beta$ and $\beta_2 = -\rho\beta$, so the agent sets $e_1 = \beta$ and $e_2 = 0$. With α chosen optimally by the principal, the principal's expected profit as a function of β is

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

and the optimal choice of β yields profit

$$\Pi^{OT}(\lambda = 1) = \frac{1}{2\delta^2\left[1 + R\left(1 - \rho^2\right)\right]}. \tag{3}$$

The SD contract induces the agent to exert effort on both tasks, while the OT contract elicits effort only on one task. However, for any given β the risk premium under the SD contract, $R\beta^2(1+\rho)$, is larger than that under the OT contract, $\frac{1}{2}R\beta^2\left(1 - \rho^2\right)$. 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, which we denote $\delta^S\left(r\sigma^2, \rho\right)$ is increasing in each of its arguments.

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 ($\lambda > 1$), there are in principle 3 possibilities: (a) both types of agent are induced to exert balanced efforts, (b) one type exerts balanced efforts and the other type focused effort, or (c) both types exert focused effort. There is also the possibility of using menus of contracts as screening devices, given the private information held by the agent. In this subsection we analyze what the principal can achieve by using optimally designed menus of deterministic contracts.

We first show that it is not in fact possible for the principal to induce both types of agent to

choose balanced efforts. Achieving this would necessitate a menu of the form

$$C_1 \quad : \quad w_1 = \alpha_1 + \beta_1 x_1 + \lambda \beta_1 x_2,$$
$$C_2 \quad : \quad w_2 = \alpha_2 + \beta_2 x_2 + \lambda \beta_2 x_1.$$

If, for each i, contract C_i were chosen by agent i, then balanced efforts would be induced from both agent types. Now note that for agent 1 to be willing to choose contract C_1 requires

$$ACE_1(C_1) \geq ACE_1(C_2),$$

where the notation ACE denotes the certainty-equivalent achieved by the agent. The analogous self-selection constraint for agent 2 is clearly

$$ACE_2(C_2) \geq ACE_2(C_1).$$

Since $\lambda > 1$ it must be that

$$ACE_2(C_1) > ACE_1(C_1), \tag{4}$$

since agent 1's certainty-equivalent from contract C_1 equals that he would obtain from focusing all his effort on task 1 (which is one of his optimal effort allocations), whereas agent 2's certainty equivalent from C_1 equals that he would obtain from focusing all his effort on task 2 (which is his unique optimal effort choice), and task 2 is more highly rewarded than task 1 in contract C_1. Similarly, for all $\lambda > 1$ we have

$$ACE_1(C_2) > ACE_2(C_2). \tag{5}$$

If $ACE_1(C_1) \geq ACE_2(C_2)$, then inequality (4) implies that $ACE_2(C_1) \geq ACE_2(C_2)$, so the self-selection constraint for agent 2 would be violated. If, instead, $ACE_1(C_1) < ACE_2(C_2)$, then inequality (5) implies that $ACE_1(C_1) < ACE_1(C_2)$, so the self-selection constraint for agent 1 would be violated. Therefore, balanced efforts by both types of agents cannot be achieved.

Furthermore, since faced with a menu of deterministic linear contracts, an agent either is willing to exert balanced efforts or strictly prefers fully focused efforts, the above argument also shows that it is not possible for the principal to induce both types of agent to exert strictly positive efforts on both tasks.

Asymmetric Deterministic Menus A menu of contracts can, however, be designed in such a way as to induce one type of agent to exert balanced efforts. The natural question is then what is the optimal menu to achieve this. It turns out that the optimal deterministic scheme (an "Asymmetric Deterministic Menu" or ADM) which induces balanced efforts from one type of agent has the following form:

$$C_1 \quad : \quad w_1 = \alpha_1 + \beta_1 x_1 - \rho \beta_1 x_2,$$
$$C_2 \quad : \quad w_2 = \alpha_2 + \beta_2 x_2 + \lambda \beta_2 x_1.$$

In this menu, C_i is the contract intended for agent i. If agent 2 does choose C_2, he would be willing to choose balanced efforts $e_1 = e_2 = \frac{\beta_2}{1+\lambda}$. If agent 1 chooses C_1, he would set $e_1 = \beta_1$ and $e_2 = 0$. The negative coefficient on x_2 in C_1 exploits the correlation in the shocks to the performance measures to improve the insurance offered to agent 1.

The certainty-equivalents that each of the two contracts offers to each of the two types of agent are:

$$
\begin{aligned}
ACE_1\left(C_1\right) &= \alpha_1 + \frac{\left(\beta_1\right)^2}{2} - \frac{1}{2} R\left(\beta_1\right)^2 \left(1 - \rho^2\right), \\
ACE_1\left(C_2\right) &= \alpha_2 + \frac{\lambda^2\left(\beta_2\right)^2}{2} - \frac{1}{2} R\left(\beta_2\right)^2 \left(\lambda^2 + 2\rho\lambda + 1\right), \\
ACE_2\left(C_1\right) &= \alpha_2 + \frac{\left(\beta_2\right)^2}{2} - \frac{1}{2} R\left(\beta_2\right)^2 \left(\lambda^2 + 2\rho\lambda + 1\right), \\
ACE_2\left(C_2\right) &= \alpha_1 + \frac{\left(\beta_1\right)^2}{2\lambda^2} - \frac{1}{2} R\left(\beta_1\right)^2 \left(1 - \rho^2\right).
\end{aligned}
$$

The problem faced by the principal is to choose $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ to maximize

$$
\frac{1}{2}\left[\frac{\beta_1}{\delta} - \alpha_1 - \left(\beta_1\right)^2\right] + \frac{1}{2}\left[\left(\frac{\beta_2}{1+\lambda}\right)\left(\frac{1+\delta}{\delta}\right) - \alpha_2 - \left(\beta_2\right)^2\right],
$$

subject to participation and self-selection constraints for both types of agent:

$$
\begin{aligned}
ACE_2\left(C_2\right) &\geq 0, \\
ACE_2\left(C_2\right) &\geq ACE_2\left(C_1\right), \\
ACE_1\left(C_1\right) &\geq 0, \text{ and} \\
ACE_1\left(C_1\right) &\geq ACE_1\left(C_2\right).
\end{aligned}
$$

Since for all $\lambda > 1$ we have $ACE_1\left(C_2\right) \geq ACE_2\left(C_2\right)$, the second and fourth constraints above imply that the third constraint will not bind, and hence agent 1 earns an "information rent".

For the two self-selection constraints to be satisfied simultaneously, it is necessary that

$$
\frac{\left(\beta_2\right)^2}{2} - \frac{\left(\beta_1\right)^2}{2\lambda^2} \geq \frac{\lambda^2\left(\beta_2\right)^2}{2} - \frac{\left(\beta_1\right)^2}{2},
$$

which is equivalent to

$$
\beta_1 \geq \lambda\beta_2.
$$

For given (β_1, β_2), it is optimal for the principal to set α_2 so agent 2's participation constraint binds and to set α_1 so agent 1's self-selection constraint binds. Then the constraint $\beta_1 \geq \lambda\beta_2$ is both necessary and sufficient for agent 2 to be willing to choose C_2. We may then restate the principal's

problem as

$$\max_{\beta_1,\beta_2} \left\{ \begin{array}{l} \frac{1}{2}\left[\frac{\beta_1}{\delta} - \frac{(\beta_1)^2}{2} - \frac{1}{2}R(\beta_1)^2\left(1-\rho^2\right) - \left(\lambda^2-1\right)\frac{(\beta_2)^2}{2}\right] \\ +\frac{1}{2}\left[\left(\frac{\beta_2}{1+\lambda}\right)\left(\frac{1+\delta}{\delta}\right) - \frac{(\beta_2)^2}{2} - \frac{1}{2}R(\beta_2)^2\left(\lambda^2+2\rho\lambda+1\right)\right] \end{array} \right\}$$

subject to

$$\beta_1 \geq \lambda\beta_2.$$

If we initially ignore the constraint, the first-order conditions with respect to β_1 and β_2 yield

$$\beta_1^* = \frac{1}{\delta\left(1+R\left(1-\rho^2\right)\right)},$$

$$\beta_2^* = \frac{1+\delta}{\delta\left(1+\delta\right)\left(\lambda^2 R\left(\lambda^2+2\rho\lambda+1\right)\right)}.$$

These values satisfy the constraint $\beta_1^* \geq \lambda\beta_2^*$ if and only if

$$\left(\frac{\lambda^2+R\left(\lambda^2+2\rho\lambda+1\right)}{1+R\left(1-\rho^2\right)}\right)\left(\frac{\lambda}{1+\lambda}\right) \geq 1+\delta. \tag{6}$$

Since the left-hand-side of this inequality is increasing in R, λ, and ρ, larger values of these three variables make it easier to satisfy the constraint. But, so long as R is finite, for δ sufficiently large it must be that $\beta_1^* < \lambda\beta_2^*$. Since the objective function is concave in β_1 and β_2 and the constraint set is linear in these arguments, any unconstrained solution which violates the constraint must involve the constraint binding at the optimum.

Thus when (6) is violated, the principal's problem is equivalent to

$$\max_{\beta_2}\left\{\beta_2\left[\frac{\left(\lambda^2+\lambda+\delta+1\right)}{2\delta\left(1+\lambda\right)}\right] - \frac{\lambda^2\left(\beta_2\right)^2}{2} - \frac{1}{2}R\left(\beta_2\right)^2\left[\left(1-\frac{\rho^2}{2}\right)\lambda^2+\rho\lambda+\frac{1}{2}\right]\right\}.$$

With β_2 chosen optimally, the principal's profit from this "constrained" ADM (ADMC) is

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

On the other hand, when (6) is satisfied, so the optimal ADM is "unconstrained" (ADMU), the principal's profit is

$$\Pi^{ADMU} = \frac{1}{4\delta^2}\left[\frac{1}{1+R\left(1-\rho^2\right)} + \frac{(1+\delta)^2}{(1+\lambda)^2}\left(\frac{1}{\lambda^2+R\left(\lambda^2+2\rho\lambda+1\right)}\right)\right].$$

12

Symmetric Deterministic Menus   In contrast to an ADM, a menu of contracts of the following form induces both types of agents to focus their effort on their preferred task:

$$C_1 \; : \; w_1 = \alpha + \beta x_1 - \rho \beta x_2,$$
$$C_2 \; : \; w_2 = \alpha + \beta x_2 - \rho \beta x_1,$$

where contract C_i will be chosen by agent i. C_1 induces agent 1 to choose $e_1 = \beta$ and $e_2 = 0$, while C_2 induces agent 2 to choose $e_1 = 0$ and $e_2 = \beta$. The negative coefficient in each contract on the agent's less-preferred task improves insurance by exploiting the correlation in the shocks.   The certainty equivalent for both types of agent is

$$ACE = \alpha + \frac{\beta^2}{2} - \frac{1}{2} R \beta^2 \left(1 - \rho^2 \right).$$

With α chosen optimally, the principal's problem is then simply

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

Hence maximized profit from the SDM is

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

which is the same as the maximal profit under the OT contract when $\lambda = 1$. Profit under the SDM is independent of λ, and for all $\lambda > 1$ and $\rho \geq 0$,

$$\Pi^{SDM} > \Pi^{SD}.$$

The Optimal Deterministic Scheme   In light of the preceeding analysis, when $\lambda > 1$, the candidate optima are the asymmetric deterministic menu and the symmetric deterministic menu. Under the ADM, providing incentives for one type of agent to choose balanced efforts imposes more risk on that agent than he would bear under the SDM.   Thus, which of these is optimal depends on the benefit to the principal from inducing balanced efforts relative to the cost of imposing the required risk.   This benefit is determined by the degree of complementarity between the tasks in the principal's benefit function For any $\lambda > 1$, there exists a critical value of δ, denoted $\delta^D \left(\lambda, r\sigma^2, \rho \right)$ above which ADM dominates SDM and below which the opposite is true.   We formalize this as:

Proposition 1 *There exist values of δ, $\delta^{SDM/ADMU} \leq \delta^{ADMU/ADMC}$, such that for $\delta < \delta^{SDM/ADMU}$, a SDM is the optimal deterministic scheme, for $\delta \in \left(\delta^{SDM/ADMU}, \delta^{ADMU/ADMC} \right)$, the optimal deterministic scheme is an ADM and $\beta_1 > \lambda \beta_2$, and for $\delta > \delta^{ADMU/ADMC}$, the optimal deterministic scheme is an ADM and $\beta_1 = \lambda \beta_2$. Moreover, for all $\lambda > 1$, $\delta^{SDM/ADMU} > \lambda$, $\delta^{SDM/ADMU} - \lambda$ is increasing in λ, and $\delta^{ADMU/ADMC} - \lambda$ in increasing in λ.*

We can summarize the conclusions from our analysis of the optimal deterministic schemes graphically, as follows.

At $\lambda=1$, SD is optimal for $\delta > \delta^S(r\sigma^2,\rho)$ and SDM for $\delta < \delta^S(r\sigma^2,\rho)$.
For $\lambda > 1$, ADM is optimal for $\delta > \delta^D(\lambda,r\sigma^2,\rho)$ and SDM for $\delta < \delta^D(\lambda,r\sigma^2,\rho)$.
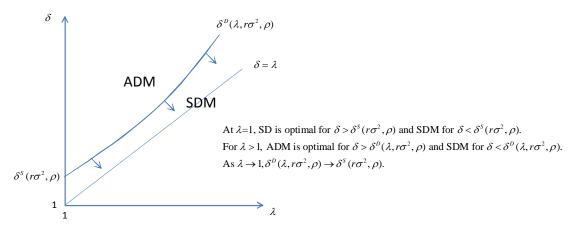As $\lambda \to 1, \delta^D(\lambda,r\sigma^2,\rho) \to \delta^S(r\sigma^2,\rho)$.

Figure 1: Optimal Deterministic Schemes

3.2 Random Contracts

3.2.1 Ex Ante Randomization

The simplest contract involving *ex ante randomization* specifies that with probability $\frac{1}{2}$, $\beta_1 = \beta$ and $\beta_2 = 0$ and with probability $\frac{1}{2}$, $\beta_1 = 0$ and $\beta_2 = \beta$. Under this contract, which we term EAR, the principal commits to employ a randomizing device to determine on which of the two outputs the agent's pay will be based.

Proposition 2 *1. Let $\overline{e}^{EAR}$ 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 induces interior solutions for the agents' effort choices, both types of agent choose $(\overline{e}^{EAR}, \underline{e}^{EAR})$ satisfying*

$$\overline{e}^{EAR} + \lambda \underline{e}^{EAR} = \frac{\beta}{\lambda + 1} \tag{7}$$

$$\lambda = \exp\left[r\beta(\overline{e}^{EAR} - \underline{e}^{EAR})\right]. \tag{8}$$

Thus

$$\overline{e}^{EAR} = \frac{\beta}{(\lambda + 1)^2} + \frac{\lambda \ln \lambda}{r\beta(\lambda + 1)}$$

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

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

2. The principal's profit from interior effort choices by the agents under ex ante randomization, for

14

a given β, is

$$\Pi^{EAR}(\beta) \ = \ \underline{e}^{EAR} + \frac{1}{\delta}\overline{e}^{EAR} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2}R\beta^2 - \frac{\ln\left(\frac{(\lambda+1)^2}{4\lambda}\right)}{2r} \tag{9}$$

$$= \ \frac{(\delta+1)\beta}{\delta\,(\lambda+1)^2} - \frac{(\delta-\lambda)\ln\lambda}{\delta r\beta\,(\lambda+1)} - \frac{\beta^2}{2\,(\lambda+1)^2} - \frac{R\beta^2}{2} - \frac{\ln\left(\frac{(\lambda+1)^2}{4\lambda}\right)}{2r}. \tag{10}$$

To understand equations (7) and (8), 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 $\beta = \partial c/\partial \overline{e} + \partial c/\partial \underline{e}$ for both types of agent, which gives us equation (7). Furthermore, for both types of agent, $\frac{\partial c/\partial \underline{e}}{\partial c/\partial \overline{e}} = \lambda$, and the first-order conditions for interior optimal efforts imply

$$\lambda = \frac{\frac{\partial c}{\partial \underline{e}}}{\frac{\partial c}{\partial \overline{e}}} = \frac{E\left[U'(\cdot)I_{\{\underline{x} \text{ is rewarded}\}}\right]}{E\left[U'(\cdot)I_{\{\overline{x} \text{ is rewarded}\}}\right]} = \exp\left[r\beta(\overline{e} - \underline{e})\right], \tag{11}$$

which gives us equation (8).

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 (8) shows, optimal self-insurance will result in a smaller gap between the effort levels on the two tasks, the smaller the cost difference between tasks (smaller λ), the more risk-averse is the agent (larger r), and the higher the incentive intensity (larger β).

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 ex ante randomization is more robust to the introduction of private information on the part of the agent than is a symmetric deterministic contract. EAR is also more robust to uncertainty about the magnitude of λ than is an asymmetric deterministic menu (ADM). If the principal tries to design an ADM 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.

Under EAR, 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 would a deterministic contract which based pay only on a single task at rate β. This can be seen in the principal's profit expression (9): the cost of the risk imposed on the agent under ex ante random-

ization is $\frac{1}{2}R\beta^2 + \frac{1}{2r}\ln\left(\frac{(\lambda+1)^2}{4\lambda}\right)$, whereas under a deterministic contract that based pay on a single task at rate β, it would be only $\frac{1}{2}R\beta^2$.

Remark 1 *We have established Proposition 2 under the assumption that the principal can commit to randomizing half-half between the two compensation formulae.[8] It is natural to wonder whether the same outcome would result if, instead, the principal chooses the randomizing probability at the same time as the agent chooses efforts (we term this "interim randomization"). We can prove that under interim randomization, the unique Bayes-Nash equilibrium is the same as the outcome described in Proposition 2.[9] Thus the attractive properties of ex ante randomization are not crucially dependent on the principal's having the power to commit to the randomizing probability.*

3.2.2 Generalized Ex Ante Randomization

Under a more general form of ex ante randomization, which we will term gEAR, the contract specifies that with probability $\frac{1}{2}$, the agent will be compensated according to $w = \alpha + \beta x_1 + k\beta x_2$, and with probability $\frac{1}{2}$, he will be compensated according to $w = \alpha + \beta x_2 + k\beta x_1$, where the parameter $k \in (-1,1)$. EAR is just the special case of gEAR where $k = 0$. gEAR makes the agent uncertain at the time he chooses his efforts about which performance measure will be more highly rewarded, and by varying the level of k, the principal can affect how much this uncertainty matters to the agent. The closer k is to 1, the more similar are the two possible compensation schedules, so intuitively, the weaker are the agent's incentives to choose balanced efforts as a means of insuring himself against the risk stemming from the principal's randomization. If k were set equal to 1, the randomness in the compensation scheme would completely disappear, and the contract would collapse to the symmetric deterministic (SD) scheme.

Proposition 3 *1. Under gEAR, $k < \frac{1}{\lambda}$ is a necessary condition for each agent's optimal efforts on both tasks to be strictly positive.*
2. For each agent, let $\overline{e}^{gEAR}$ denote the effort exerted on his less-costly task and $\underline{e}^{gEAR}$ the effort on his more-costly task. Then for a given $k \in (-1, \frac{1}{\lambda})$, if gEAR induces interior solutions for the

---

[8] Given the power to commit to a randomizing probability, it is optimal for the principal to commit to rewarding each of the two tasks with equal probability. This 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.

[9] To see that the outcome described in Proposition 2 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 reward. Given that the principal uses randomizing probability $\frac{1}{2}$, 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 to compensate task 1 with $p > 1/2$ ($p < 1/2$), 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 compensate task 2 (task 1).

agents' effort choices, both types of agent choose $(\overline{e}^{gEAR}, \underline{e}^{gEAR})$ *satisfying*

$$\overline{e}^{gEAR} + \lambda\underline{e}^{gEAR} = \frac{\beta(1+k)}{\lambda+1} \tag{12}$$

$$\overline{e}^{gEAR} - \underline{e}^{gEAR} = \frac{\ln\frac{\lambda-k}{1-k\lambda}}{r\beta(1-k)}. \tag{13}$$

3. *The gap in efforts,* $\overline{e}^{gEAR} - \underline{e}^{gEAR}$, *is increasing in* λ, *approaching 0 as* $\lambda \to 1$; *decreasing in* r, *approaching 0 as* $r \to \infty$; *and increasing in* k, *approaching 0 as* $k \to -1^+$.

4. *The principal's profit from interior effort choices by the agents under gEAR, for given* $\beta > 0$ *and* $k \in (-1, \frac{1}{\lambda})$, *is*

$$\Pi^{gEAR}(\beta, k) = \underline{e}^{gEAR} + \frac{1}{\delta}\overline{e}^{gEAR} \quad - \quad \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} \tag{14}$$
$$- \quad \frac{1}{2}R\beta^2(1+2\rho k + k^2) - \frac{1}{2r}\ln\left(\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right)$$

To understand the first part of the proposition, note that if $k \geq \frac{1}{\lambda}$, then *whichever of the two compensation schemes is randomly selected*, the ratio of marginal return to marginal cost is at least as large for effort on the preferred task as for effort on the less-preferred task, for both types of agent and for any pair of effort levels. Hence in this case, both types of agent would optimally exert effort only on their preferred task.

Equation (12) shows, not surprisingly, that "aggregate effort" $\overline{e}^{gEAR} + \lambda\underline{e}^{gEAR}$ is increasing in k. More interesting, equation (13) shows that as k increases, the gap in efforts, $\overline{e}^{gEAR} - \underline{e}^{gEAR}$, also increases. The larger is k, the more similar are the two possible compensation schedules, so the less risk the randomization per se imposes on the agent and hence the weaker are his incentives to self-insure by choosing relatively balanced efforts. As k approaches -1, the self-insurance motive approaches its strongest level, and the optimal gap in efforts approaches 0. Just as with simple EAR, the optimal gap in efforts is increasing in λ, measuring the strength of the agent's preference across tasks, and decreasing in the agent's risk aversion, r.

In the principal's profit expression (14), the final two terms (those on the second line) represent the total cost of the risk borne by the agent. The first of these terms is the risk cost that would be imposed by a deterministic contract of the form $w = \alpha + \beta x_1 + k\beta x_2$ (or equivalently, $w = \alpha + \beta x_2 + k\beta x_1$), and the second is the additional risk cost imposed by the principal's randomization. As long as the agent's optimally chosen degree of self-insurance is only partial, the total risk costs imposed by gEAR exceed those imposed by a deterministic contract corresponding to the same values of β and k.

To understand the effect of varying the parameter k on the principal's profit from gEAR, it is helpful to define the variable $B \equiv \beta(1+k)$, because aggregate effort $\overline{e} + \lambda\underline{e}$ is proportional to B.

The principal's profit expression (14) can then be re-expressed as a function of B and k:

$$\Pi^{gEAR}(B,k) = \frac{(\delta+1)B}{\delta\left(\lambda+1\right)^2} - \frac{(\delta-\lambda)\ln\left(\frac{\lambda-k}{1-k\lambda}\right)}{\delta(\lambda+1)rB\left(\frac{1-k}{1+k}\right)} - \frac{B^2}{2\left(\lambda+1\right)^2}$$

$$- \frac{1}{2}RB^2\left(\frac{1+2\rho k+k^2}{(1+k)^2}\right) - \frac{1}{2r}\ln\left(\frac{(\lambda+1)^2\left(1-k\right)^2}{4(1-k\lambda)(\lambda-k)}\right). \quad (15)$$

Holding B fixed and varying k allows us to identify the effect of k on the principal's profit from inducing any given level of aggregate effort. Equation (15) shows that increasing k has three effects. First, a larger k raises the gap between efforts on the two tasks and, with B and hence aggregate effort held fixed, this larger gap translates into a lower benefit for the principal whenever $\delta > \lambda$, i.e. whenever the principal's desire for balanced efforts is stronger than the agent's preference across tasks. As a result, the principal's profit falls, as reflected in the effect of k on the second term in equation (15). Second, a larger k, by inducing the agent to choose less balanced efforts, actually results in the agent bearing more risk from the randomization per se and hence reduces the principal's profit; this effect is reflected in the fifth term in equation (15). Finally, a larger k reduces the cost (per unit of aggregate effort induced) of the risk imposed on the agent from the shocks to measured performance. This improved diversification raises the principal's profit, as reflected in the fourth term in equation (15).

In general, therefore, the optimal design of a generalized ex ante randomization scheme involves a trade-off between these three different effects. Weighting the different performance measures more equally in the two possible compensation schedules is costly in terms of effort balance and thereby in terms of the risk cost that randomization imposes on the agent, but it is helpful in allowing better diversification in the face of the random shocks to measured performance. In Section 5, where we identify environments where randomized schemes outperform deterministic ones, we will explicitly study how the optimal choice of k varies across these environments. For now, though, we turn to a second class of randomized contracts.

3.2.3   Ex Post Randomization

Under the simplest 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 simple ex ante randomization, the agent is uncertain about which performance indicator will determine his pay, but with ex post randomization (EPR), unlike with EAR, 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 4 *1. When ex post randomization induces interior solutions for the agents' effort choices, each type of agent chooses effort on his less-costly task, $\overline{e}^{EPR}$, and effort on his more-costly task, $\underline{e}^{EPR}$, satisfying*

$$\overline{e}^{EPR} + \lambda \underline{e}^{EPR} = \frac{\beta}{\lambda + 1} \tag{16}$$

$$\lambda = \exp\left[r\beta(\overline{e}^{EPR} - \underline{e}^{EPR})\right] \frac{\Phi\left(\frac{\overline{e}^{EPR} - \underline{e}^{EPR} + R\beta(1-\rho)}{\theta}\right)}{\Phi\left(\frac{-(\overline{e}^{EPR} - \underline{e}^{EPR}) + R\beta(1-\rho)}{\theta}\right)}, \tag{17}$$

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

$$\begin{aligned}
\Pi^{EPR}(\beta) = {} & \underline{e}^{EPR} + \frac{1}{\delta}\overline{e}^{EPR} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2}R\beta^2 \\
& - \frac{1}{r}\ln\left[\frac{1+\lambda}{\lambda}\Phi\left(\frac{\overline{e}^{EPR} - \underline{e}^{EPR} + R\beta(1-\rho)}{\theta}\right)\right] \\
& - \beta(\overline{e}^{EPR} - \underline{e}^{EPR})\Phi\left(-\frac{(\overline{e}^{EPR} - \underline{e}^{EPR})}{\theta}\right) + \beta\theta\phi\left(\frac{\overline{e}^{EPR} - \underline{e}^{EPR}}{\theta}\right),
\end{aligned} \tag{18}$$

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

Proposition 4 shows that aggregate effort, $\overline{e} + \lambda\underline{e}$, is the same under simple EPR as under simple EAR—compare equations (16) and (7). 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, $(\lambda + 1)(\overline{e} + \lambda\underline{e})$. Just as for EAR, the first-order conditions for interior optimal efforts imply

$$\lambda = \frac{\frac{\partial c}{\partial \underline{e}}}{\frac{\partial c}{\partial \overline{e}}} = \frac{E\left[U'(\cdot)I_{\{\underline{x} \text{ is rewarded}\}}\right]}{E\left[U'(\cdot)I_{\{\overline{x} \text{ is rewarded}\}}\right]}, \tag{19}$$

but for ex post randomization

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

which when combined with equation (19) gives us equation (17).

Under EAR, 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

19

choice of which task to reward. Under EPR, 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 $\overline{e} - \underline{e}$ 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 $\overline{e} - \underline{e}$. Formally, the right-hand side of equation (17), which is increasing in $\overline{e} - \underline{e}$, is strictly greater than the right-hand side of equation (8) for all $\overline{e} - \underline{e} > 0$, so

$$\overline{e}^{EPR} - \underline{e}^{EPR} < \overline{e}^{EAR} - \underline{e}^{EAR} \qquad \forall \lambda > 1. \tag{20}$$

Equation (17) also allows us to show that under EPR, 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 EAR. Furthermore, while σ^2 and ρ have no effect on the gap in effort levels under EAR, under EPR 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 $\overline{e} - \underline{e}$ 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 in Section 5.)

Intuitively, we would expect that the principal's freedom, under EPR, to choose the performance measure that minimizes her wage bill would result in weaker overall incentives for the agent than under EAR. This intuition is correct in the sense that the *sum* of the efforts on the two tasks, $\overline{e} + \underline{e}$, is lower under EPR than under EAR. 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 inequality (20) and the fact that $\overline{e} + \lambda\underline{e}$ 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 her preference for balanced efforts, is larger or smaller than λ, measuring the strength of the agents' preference across tasks.

We can also show that EPR with coefficient β, in contrast to EAR, imposes *lower* risk costs on the agent than would a deterministic contract which based pay only on a single task at rate β. Formally, this claim corresponds to the result that the principal's profit, given by equation (18), is *greater* than $\underline{e}^{EPR} + \frac{1}{\delta}\overline{e}^{EPR} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2}R\beta^2$. The intuitive reason why this result holds is that the variance of the wage under EPR, $w = \min\{\alpha + \beta x_1, \alpha + \beta x_2\}$, is lower than the variance of either $\alpha + \beta x_1$ or $\alpha + \beta x_2$.[10][11]

---

[10] We can also show that the risk costs imposed on the agent by EPR are increasing in the gap $\overline{e} - \underline{e}$, reflecting the fact that the variance of $w = \min\{\alpha + \beta x_1, \alpha + \beta x_2\}$ is increasing in $\overline{e} - \underline{e}$.

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

3.2.4 Generalized Ex Post Randomization

A contract involving ex post randomization, like one involving ex ante randomization, can be generalized to allow arbitrary relative weights on the two performance measures. Specifically, under a generalized ex post randomization (gEPR) scheme, the principal, after observing x_1 and x_2, chooses whether to pay the agent $\alpha + \beta x_1 + k\beta x_2$ or $\alpha + \beta x_2 + k\beta x_1$, where again the parameter $k \in (-1, 1)$. As with gEAR, the closer k is to 1, the more similar are the two possible compensation formulae from which the principal chooses, and if k were actually set equal to 1, the scheme would involve no randomness at all and would collapse to the SD scheme.

Proposition 5 *1. Under gEPR, $k < \frac{1}{\lambda}$ is a necessary condition for each agent's optimal efforts on both tasks to be strictly positive.*
2. When for a given $k \in (-1, \frac{1}{\lambda})$, gEPR induces interior solutions for the agents' effort choices, each type of agent chooses effort on his less-costly task, $\overline{e}^{gEPR}$, and effort on his more-costly task, $\underline{e}^{gEPR}$, satisfying

$$\overline{e}^{gEPR} + \lambda\underline{e}^{gEPR} \;=\; \frac{\beta(1+k)}{\lambda+1} \tag{21}$$

$$\frac{\lambda - k}{1 - k\lambda} \;=\; \exp\left[r\beta(1-k)(\overline{e}^{gEPR} - \underline{e}^{gEPR})\right] \frac{\Phi\left(\frac{(1-k)(\overline{e}^{gEPR} - \underline{e}^{gEPR}) + r(\sigma^g)^2\beta(1-\rho^g)}{\theta^g}\right)}{\Phi\left(\frac{-(1-k)(\overline{e}^{gEPR} - \underline{e}^{gEPR}) + r(\sigma^g)^2\beta(1-\rho^g)}{\theta^g}\right)} \tag{22}$$

where $(\sigma^g)^2 \equiv var(x_1 + kx_2) = \sigma^2(1 + 2\rho k + k^2)$, $\rho^g \equiv corr(x_1 + kx_2, x_2 + kx_1) = \frac{\rho + 2k + \rho k^2}{1 + 2\rho k + k^2}$, and $\theta^g \equiv \sigma^g[2(1 - \rho^g)]^{\frac{1}{2}}$.
3. When for a given $\lambda > 1$ and $k \in (-1, \frac{1}{\lambda})$, gEPR and gEAR both induce interior solutions for efforts, then $\overline{e}^{gEPR} - \underline{e}^{gEPR} < \overline{e}^{gEAR} - \underline{e}^{gEAR}$. 4. The principal's profit from interior effort choices by the agents under gEPR, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$, is

$$\begin{aligned}
\Pi^{gEPR}(\beta, k) \;=\;\; & \underline{e}^{gEPR} + \frac{1}{\delta}\overline{e}^{gEPR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^g)^2\beta^2 \tag{23} \\
& - \frac{1}{r}\ln\left[\exp\{-r\beta(1-k)(\overline{e}^{gEPR} - \underline{e}^{gEPR})\}\Phi(-) + \Phi(+)\right] \\
& - \beta(1-k)(\overline{e}^{gEPR} - \underline{e}^{gEPR})\Phi\left(-\frac{(1-k)(\overline{e}^{gEPR} - \underline{e}^{gEPR})}{\theta^g}\right) \\
& + \beta\theta^g\phi\left(\frac{(1-k)(\overline{e}^{gEPR} - \underline{e}^{gEPR})}{\theta^g}\right).
\end{aligned}$$

where

$$\begin{aligned}
\Phi(-) \;\equiv\;\; & \Phi\left(\frac{-(1-k)(\overline{e}^{gEPR} - \underline{e}^{gEPR}) + r(\sigma^g)^2\beta(1-\rho^g)}{\theta^g}\right) \\
\Phi(+) \;\equiv\;\; & \Phi\left(\frac{(1-k)(\overline{e}^{gEPR} - \underline{e}^{gEPR}) + r(\sigma^g)^2\beta(1-\rho^g)}{\theta^g}\right)
\end{aligned}$$

Comparison of equations (12) and (21) shows that for any $k \in (-1, \frac{1}{\lambda})$ that induces interior effort choices under both gEPR and gEAR, aggregate effort $\overline{e} + \lambda \underline{e}$ is equal under the two forms of randomization. As a consequence, the agent's total cost of efforts incurred is also equal under the two schemes. On the other hand, comparison of equations (13) and (22) shows that (as part 3 of Proposition 5 summarizes), gEPR induces a smaller gap in efforts on the two tasks than gEAR, for any given k. Under gEPR, not only does partially balancing efforts provide the agent with some insurance against the risk stemming from the (ex ante) randomness of the compensation schedule, as is the case under gEAR, but also, under gEPR, the principal's strategic ex post choice of which compensation schedule to employ means that the more the agent focuses his effort on his preferred task, the lower the relative marginal return to that task.

With respect to the risk costs imposed on the agent by gEPR, we can show that for any given β and k, gEPR imposes *lower* risk costs than would either of the deterministic contracts $w = \alpha + \beta x_1 + k\beta x_2$ or $w = \alpha + \beta x_2 + k\beta x_1$. This lower risk reflects the fact that the variance of the wage under gEPR, $w = \min\{\alpha + \beta x_1 + k\beta x_2, \alpha + \beta x_2 + k\beta x_1\}$, is lower than the variance of either $\alpha + \beta x_1 + k\beta x_2$ or $\alpha + \beta x_2 + k\beta x_1$. Section 3.2.2 showed, by contrast, that for any given β and k, gEAR imposes *higher* risk costs than would either of the deterministic contracts above.

3.2.5 Generalized Ex Ante versus Generalized Ex Post Randomization

The preceding paragraphs have argued that (assuming interior solutions for the agents' efforts), for any given β and k, i)gEPR induces a strictly smaller gap in efforts $\overline{e} - \underline{e}$ than gEAR, while the two schemes induce the same aggregate effort $\overline{e} + \lambda \underline{e}$ and hence the same total cost of effort, and ii) gEPR imposes lower risk costs on the agent than gEAR. Taken together, these findings generate the following proposition:

Proposition 6 *If, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$, both gEAR and gEPR induce interior solutions for the agents' effort choices, and if $\delta \geq \lambda$, then gEPR generates at least as great a profit for the principal as gEAR.*

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, generates a higher expected benefit for the principal.

4   When Are Deterministic Contracts Optimal?

This section identifies three environments in which randomized contracts are dominated by 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 across tasks. In each of these three environments, we can show that randomized contracts impose too much risk on

the agent, relative to the effort benefits they generate, and as a consequence are dominated by a symmetric deterministic scheme.

Proposition 7 *For any given $\beta > 0$ and $k \in (-1, 1)$, both gEAR and gEPR yield lower profit for the principal than a suitably designed symmetric deterministic (SD) scheme, if any of the following conditions holds:*

1. *$\lambda = 1$;*

2. *gEAR and gEPR induce the agent to exert effort only on his preferred task;*

3. *$\delta \leq \lambda$.*

The key to understanding this proposition is the finding that, for any $k \in (-1, 1)$ and for any λ, a SD scheme can induce any given level of *aggregate effort* $\overline{e} + \lambda \underline{e}$ while imposing lower risk costs on the agent than gEAR or gEPR. At the same time, we know from Section 3 that whenever $\lambda > 1$, a SD scheme always induces the agent to exert effort only on his preferred task, whereas gEAR and gEPR have the potential to induce better-balanced efforts. In general, therefore, the principal faces a trade-off in choosing between randomized and deterministic incentive schemes. Randomized schemes are typically better at inducing balanced efforts, while deterministic schemes have the advantage of imposing lower risk costs on the agent per unit of aggregate effort induced. The three conditions identified in Proposition 7 are ones under which this trade-off does not in fact arise. Under condition 1 or 2, randomized schemes are no better than a SD scheme at inducing balanced efforts: in the former case, the SD scheme, like the randomized schemes, induces perfectly balanced efforts, and in the latter case, even the randomized schemes induce corner solutions. Under condition 3, for any given level of aggregate effort, a shift towards more balanced efforts would actually reduce social surplus (at least weakly), because the benefit to the principal would be outweighed by the disutility suffered by the agent. Therefore, under any of conditions 1, 2, or 3, the potential benefits of randomized schemes in inducing better-balanced efforts do not actually materialize, and the principal's optimal incentive scheme is a deterministic one.

Proposition 7 has an informative corollary:

Corollary 1 *Consider the limiting case where $r \rightarrow 0$ and $\sigma^2 \rightarrow \infty$ in such a way that $r\sigma^2 \rightarrow R < \infty$. In this limiting case, for any $\beta > 0$ and any $k \in (-1, 1)$, 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 extremely weak incentives to choose balanced efforts.

5   When Are Random Contracts Optimal?

We now identify three environments in which random contracts, when designed optimally, can be shown to dominate the best deterministic scheme. In each of these environments, gEAR and gEPR, with the parameter k adjusted optimally, both induce the agent to choose *perfectly* balanced efforts, and gEAR is as profitable for the principal as gEPR. 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 $\lambda \to 1^+$. The second such setting is the limiting case where r goes to ∞ and σ^2 goes to 0. The final setting is that where the shocks affecting measured performance on the two tasks become perfectly correlated: $\rho \to 1$. In all three environments, we show that there is a critical degree of complementarity of tasks for the principal (i.e. a critical value of δ) above which the optimal incentive scheme is gEAR or gEPR with k adjusted optimally and below which the optimal scheme is the symmetric deterministic menu (SDM). An asymmetric deterministic menu (ADM) is never strictly optimal in these settings.

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

Consider first a setting in which the agent has private information about his preferences but the magnitude of his preference across tasks is arbitrarily small. Formally, this is the case in which $\lambda > 1$ but arbitrarily close to 1, which we term the limiting case as $\lambda \to 1^+$.

As we saw in Section 3.1.2, for any λ strictly greater than 1, the optimal deterministic scheme is either a symmetric deterministic menu (SDM) or an asymmetric deterministic menu (ADM). The latter induces one of the two types of agent to choose perfectly balanced efforts, while the former induces both types to choose fully focused efforts. Since under an ADM, providing incentives for one type to choose balanced efforts requires imposing more risk on that agent than he would bear under a SDM, the choice between an ADM and a SDM involves a trade-off for the principal between the benefits of balanced efforts and the cost of imposing risk on the agent. As we saw, for any $\lambda > 1$, there is a critical value of δ, $\delta^D(\lambda, r\sigma^2, \rho)$, above which the optimal deterministic scheme is an ADM and below which it is a SDM.

We also saw, in Section 3.1.1, that even when $\lambda = 1$, so the agent finds the two tasks equally costly, the principal faces a qualitatively similar trade-off between the benefits of balanced efforts and the cost of imposing risk on the agent. A symmetric deterministic (SD) contract that rewards the two tasks equally induces balanced efforts but imposes more risk on the agent than would a OT

24

contract that rewarded only one task and used measured performance on the other task to improve insurance. For $\lambda = 1$, there is a critical $\delta^S(r\sigma^2, \rho)$ above which the optimal deterministic scheme is a SD contract and below which it is a OT contract. This critical $\delta^S(r\sigma^2, \rho)$ is the limiting value of $\delta^D(\lambda, r\sigma^2, \rho)$ as $\lambda \to 1$. Importantly, for any $\delta > \delta^S(r\sigma^2, \rho)$, as λ increases from 1, the principal's maximum achievable profit from deterministic schemes drops discontinuously, from that achievable under a SD contract (at $\lambda = 1$) to that achievable under an ADM (for λ slightly greater than 1).

In contrast to the SD contract, under both generalized ex ante and generalized ex post randomization, the agent's optimal efforts and hence the principal's profit are continuous at $\lambda = 1$ for any value of $k \in (-1, 1)$. In the limit as $\lambda \to 1$, for both randomized schemes and for any $k \in (-1, 1)$, the gap in efforts, $\overline{e} - \underline{e}$, approaches 0. We now show that, for $\lambda > 1$ but arbitrarily close to 1, the principal can, by adjusting the parameter k optimally, achieve a profit under both gEAR and gEPR that is arbitrarily close to the profit achievable from a SD contract when $\lambda = 1$. As a consequence, for any $\delta > \delta^S(r\sigma^2, \rho)$, in the limit as $\lambda \to 1^+$, optimally designed gEAR and gEPR schemes outperform any deterministic scheme: they outperform the SD contract because that contract induces fully focused efforts whenever λ is even slightly greater than 1, and they outperform the ADM because, even for λ arbitrarily close to 1, that scheme's profit is discontinuously less than the profit from a SD contract at $\lambda = 1$.

Consider first generalized ex ante randomization. Proposition 3 established that as $\lambda \to 1$, $\overline{e}^{gEAR} - \underline{e}^{gEAR} \to 0$ for any $k \in (-1, 1)$. Equation (15) shows how varying k affects the principal's profit from gEAR, holding fixed the level of aggregate effort induced. Whereas in general, as discussed in Section 3.2.2, increasing k has opposing effects on the principal's profit, in the limit as $\lambda \to 1$, the situation is dramatically simpler. As $\lambda \to 1$, equation (15) becomes

$$\Pi^{gEAR}(B, k) = \frac{(\delta + 1)}{\delta} \frac{B}{4} - \frac{B^2}{8} - \frac{1}{2} r\sigma^2 B^2 \left(\frac{1 + 2\rho k + k^2}{(1 + k)^2} \right). \tag{24}$$

With perfectly balanced efforts ensured by λ approaching 1, an increase in k has only one effect on the principal's profit from inducing any given level of aggregate effort: it reduces the cost of the risk imposed on the agent from the shocks to measured performance, as reflected in the final term of equation (24). Hence, as $\lambda \to 1$, the principal's profit from gEAR is increasing in k as long as k induces interior solutions, which is guaranteed as long as $k < \frac{1}{\lambda}$. Therefore, as $\lambda \to 1$, the principal's profit from gEAR is maximized, for any level of aggregate effort induced, by setting k arbitrarily close to, but less than, 1. With k set in this way, the principal's profit approaches

$$\Pi^{gEAR}(B) = \frac{(\delta + 1)}{\delta} \frac{B}{4} - \frac{B^2}{8} - \frac{1}{4} r\sigma^2 B^2 (1 + \rho). \tag{25}$$

With B chosen optimally, the principal's maximized profit from the optimally designed gEAR scheme is then arbitrarily close to

$$\frac{(\delta + 1)^2}{8\delta^2(1 + 2r\sigma^2(1 + \rho))}, \tag{26}$$

which is the profit the principal would achieve, at $\lambda = 1$, from a SD contract. (Recall equation (2).

25

For generalized ex post randomization, too, as $\lambda \to 1$, $\overline{e}^{gEPR} - \underline{e}^{gEPR} \to 0$ for any $k \in (-1, 1)$, as follows from the fact that the gap in efforts on the two tasks is smaller under gEPR than under gEAR. We now convert the profit expression (23) for gEPR given in Proposition 5 into one as a function of $B \equiv \beta(1 + k)$ and k, as we did for gEAR, and simplify it in the limit as $\lambda \to 1$. This yields

$$\Pi^{gEPR}(B, k) = \frac{(\delta + 1)}{\delta} \frac{B}{4} \quad - \quad \frac{B^2}{8} - \frac{1}{2} r\sigma^2 B^2 \left(\frac{1 + 2\rho k + k^2}{(1 + k)^2}\right)$$
$$+ \quad \frac{1}{r} \left\{ \frac{rB\theta^g}{1 + k} \phi(0) - \ln\left[2\Phi\left(\frac{rB\theta^g}{2(1 + k)}\right)\right] \right\} \tag{27}$$

It can be shown that this profit expression, like that for gEAR, is increasing in k. This shows that the dominant effect of an increase in k is the improved diversification of the risk from the shocks to measured performance (as reflected in the final term on the first line). This improved diversification outweighs the cost (reflected in the sum of the two terms on the second line) of the increase in the variance of $w = \min\{\alpha + \beta x_1 + k\beta x_2, \alpha + \beta x_2 + k\beta x_1\}$ as k, and hence the correlation between $x_1 + kx_2$ and $x_2 + kx_1$, increases. Therefore, as $\lambda \to 1$, the principal's profit from gEPR is maximized, for any level of aggregate effort induced, by setting k arbitrarily close to, but less than, 1. With k set in this way, the terms on the second line in (27) approach 0 since $\rho^g \equiv corr(x_1 + kx_2, x_2 + kx_1) \to 1$ and hence $\theta^g \equiv \sigma^g[2(1 - \rho^g)]^{\frac{1}{2}} \to 0$. Hence, in the limit as $\lambda \to 1$, the principal's maximized profit from gEPR with k adjusted optimally also approaches the level he would achieve, at $\lambda = 1$, from a SD contract.

Since the principal's profit from both gEAR and gEPR is continuous at $\lambda = 1$, while her profit from a SD contract is discontinuous at $\lambda = 1$, we have established:

Proposition 8 *Consider the limiting case as $\lambda \to 1^+$. Under both gEAR and gEPR, for any given level of aggregate effort, $\overline{e} + \lambda\underline{e}$, to be induced:*

1. *the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any $k \in (-1, 1)$;*

2. *the optimal value of $k \to 1^-$;*

3. *with k adjusted optimally as $\lambda \to 1^+$, profit under the randomized schemes approaches the profit achieved from the symmetric deterministic (SD) contract **at** $\lambda = 1$.*

For $\delta > \delta^S(r\sigma^2, \rho)$, gEAR and gEPR with k adjusted optimally are more profitable than the best deterministic scheme, and for $\delta < \delta^S(r\sigma^2, \rho)$, the principal's most profitable incentive scheme is a symmetric deterministic menu (SDM).

5.2 The Limiting Case where $r \to \infty$ and $\sigma^2 \to 0$

In Section 4, we considered the limiting case where $r \to 0$ and $\sigma^2 \to \infty$ in such a way that $r\sigma^2 \to R < \infty$. We found that in this limit, for any $k \in (-1, 1)$, both gEAR and gEPR induce the

agent to exert effort only on his preferred task.

In the opposite limiting case where $r \to \infty$ and $\sigma^2 \to 0$ in such a way that $r\sigma^2 \to R < \infty$, equation (13) in Proposition 3 shows that, for any $k \in (1, \frac{1}{\lambda})$, gEAR induces the agent to choose perfectly balanced efforts: $\overline{e}^{gEAR} - \underline{e}^{gEAR} = 0$. This reflects that 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 compensation schedule, by equalizing his expected measured performance on the two tasks. Since Proposition 5 showed that $\overline{e}^{gEPR} - \underline{e}^{gEPR} \leq \overline{e}^{gEAR} - \underline{e}^{gEAR}$, it follows that in this limiting case, gEPR also induces perfectly balanced efforts—the reduction in σ^2 reinforces the agent's incentives for balance. Thus 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.

In the limit as $r \to \infty$ and $\sigma^2 \to 0$, equation (15) becomes

$$\Pi^{gEAR}(B, k) = \frac{(\delta + 1)}{\delta} \frac{B}{(\lambda + 1)^2} - \frac{B^2}{2(\lambda + 1)^2} - \frac{1}{2} RB^2 \left(\frac{1 + 2\rho k + k^2}{(1 + k)^2} \right). \tag{28}$$

Exactly as was the case when $\lambda \to 1$, with perfectly balanced efforts now ensured by $r \to \infty$ and $\sigma^2 \to 0$, an increase in k has only one effect on the principal's profit from inducing any given level of aggregate effort: it improves the diversification of the risk imposed on the agent from the shocks to performance. Hence, in this limiting case, too, the principal's profit from gEAR is increasing in k as long as k induces interior solutions, which is the case for any $k < \frac{1}{\lambda}$. Therefore, it is optimal to set k arbitrarily close to, but less than, $\frac{1}{\lambda}$. With k set in this way, the principal's profit expression in (28) approaches

$$\Pi^{gEAR}(B) = \frac{(\delta + 1)}{\delta} \frac{B}{(\lambda + 1)^2} - \frac{B^2}{2(\lambda + 1)^2} - \frac{1}{2} RB^2 \left(\frac{\lambda^2 + 2\rho\lambda + 1}{(\lambda + 1)^2} \right). \tag{29}$$

This is exactly the profit the principal would obtain from inducing aggregate effort $\frac{B}{\lambda+1}$ in a setting in which *he knew which type of agent he was employing* and designed a deterministic contract to induce balanced efforts.

In this limiting case with $r \to \infty$ and $\sigma^2 \to 0$ such that $r\sigma^2 \to R$, the principal's profit under gEPR, for given B and k, approaches the same expression as under gEAR:

$$\Pi^{gEPR}(B, k) = \frac{(\delta + 1)}{\delta} \frac{B}{(\lambda + 1)^2} - \frac{B^2}{2(\lambda + 1)^2} - \frac{1}{2} RB^2 \left(\frac{1 + 2\rho k + k^2}{(1 + k)^2} \right).$$

Hence, under gEPR, too, it is optimal to set k arbitrarily close to, but below, $\frac{1}{\lambda}$, and optimally designed gEAR does as well in this limit as optimally designed gEPR.

Proposition 9 *Consider the limiting case where $r \to \infty$ and $\sigma^2 \to 0$ in such a way that $r\sigma^2 \to R < \infty$. Under both gEAR and gEPR, for any given level of aggregate effort, $\overline{e} + \lambda\underline{e}$, to be induced:*

1. the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any λ and for any $k < \frac{1}{\lambda}$;

2. *the optimal value of* $k \to \left(\frac{1}{\lambda} \right)^{-}$;

3. *with k adjusted optimally, profit under the randomized schemes approaches the profit the prin-*
   cipal would achieve if he knew the agent's type and offered a deterministic contract inducing
   balanced efforts.

There exists $\hat{\delta}(\lambda, R, \rho)$ *such that for* $\delta > \hat{\delta}(\lambda, R, \rho)$, *gEAR and gEPR with k adjusted optimally*
are more profitable than the best deterministic scheme and for $\delta < \hat{\delta}(\lambda, R, \rho)$, *the principal's most*
profitable incentive scheme is a symmetric deterministic menu (SDM).

The critical $\hat{\delta}(\lambda, R, \rho)$ *is increasing in* λ, R, *and* ρ, *approaching* $\delta^S(R, \rho)$ *as* $\lambda \to 1$ *and approach-*
ing λ *as* $R \to 0$.

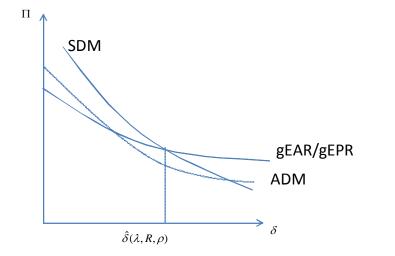The profit comparisons are summarized graphically in Figure 2.

Figure 2: Optimal incentive schemes as $r \to \infty, \sigma^2 \to 0$ s.t. $r\sigma^2 \to R < \infty$.

5.3 The Limiting Case of Perfect Correlation of the Shocks

Under generalized ex post randomization, for any $k \in (-1, 1)$, when the shocks to outputs are
perfectly correlated, the agent, given his efforts on the two tasks, faces no uncertainty about which
of $x_1 + kx_2$ or $x_2 + kx_1$ will be smaller and hence no uncertainty about whether he will be paid
$\alpha + \beta(x_1 + kx_2)$ or $\alpha + \beta(x_2 + kx_1)$. He is certain that if e_1 is greater (less) than e_2, $x_2 + kx_1$ will be
less (greater) than $x_1 + kx_2$. As a consequence, under gEPR, as long as $k < \frac{1}{\lambda}$, the agent's optimal
efforts on the two tasks will be equal. To see this, consider the agent of type 1 and suppose he
considers switching from an effort pair with $e_1 > e_2$ to one with equal efforts e^* on the two tasks,
where e^* is chosen so that aggregate effort is the same for the two effort pairs, i.e. $e_1 + \lambda e_2 = e^* + \lambda e^*$.
Then in both cases, since $\rho = 1$, the agent is certain to be paid $\alpha + \beta(x_2 + kx_1)$. Thus both the
agent's cost of effort and the risk premium are the same under the two effort pairs. The only effect

of switching from (e_1, e_2) to (e^*, e^*) on the agent's expected utility is the effect of the switch on the expected wage, which is

$$(e^* - e_2) + k(e^* - e_1) = (e^* - e_2) - k\lambda(e^* - e_2) = (e^* - e_2)(1 - k\lambda),$$

and this expression is strictly positive whenever $k < \frac{1}{\lambda}$. A switch by a type-1 agent from (e_1, e_2) with $e_1 < e_2$ to (e^*, e^*) such that aggregate effort was unchanged would also affect his expected utility only via the expected wage, which would now increase for all $k < \lambda$ (which holds by assumption). Symmetric arguments hold for a type-2 agent. Thus, when $\rho = 1$, whenever $k < \frac{1}{\lambda}$, for any given level of aggregate effort exerted, both types of agent always strictly prefer to exert equal efforts on the two tasks under gEPR.

As a consequence, in searching for *either* type of agent's optimal (e_1, e_2) for a given β and $k \in (-1, \frac{1}{\lambda})$, we can confine attention to pairs such that $e_1 = e_2 = e$. For such pairs, the expected utility of both types of agent, given $\rho = 1$, is

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

because the agent will receive a wage with the same distribution as $\alpha + \beta(1+k)e + \beta(1+k)u$, where $u \sim N(0, \sigma^2)$, that is, u has the same distribution as both of the perfectly correlated shocks.

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

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

With $\rho = 1$, and α chosen optimally by the principal, the principal's profit under gEPR for a given β and $k \in (-1, \frac{1}{\lambda})$ is

$$\Pi^{gEPR}(\beta, k) = \frac{\delta+1}{\delta}\frac{\beta(1+k)}{(\lambda+1)^2} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2(1+k)^2\beta^2. \tag{30}$$

Defining, as before, $B \equiv \beta(1+k)$, so as to examine the effect of varying k holding fixed aggregate effort, we have

$$\Pi^{gEPR}(B, k) = \frac{\delta+1}{\delta}\frac{B}{(\lambda+1)^2} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2 B^2. \tag{31}$$

The profit expression (34) is independent of k as long as $k \in (-1, \frac{1}{\lambda})$. For any $k \in (-1, \frac{1}{\lambda})$, not only are efforts perfectly balanced because $\rho = 1$, but also, because $\rho = 1$, varying k has no effect on the diversification of the risk from the shocks to performance. Under gEPR, when $\rho = 1$, any value of $k \in (-1, \frac{1}{\lambda})$ is therefore optimal. With B chosen optimally, the principal's maximized profit from gEPR in this environment is

$$\Pi^{gEPR} = \frac{(\delta+1)^2}{2\delta^2(\lambda+1)^2\left[1 + R(\lambda+1)^2\right]}. \tag{32}$$

29

Under generalized ex ante randomization, the principal's profit, expressed as a function of B and $k \in (-1, \frac{1}{\lambda})$, evaluated at $\rho = 1$, is

$$\Pi^{gEAR}(B,k) = \frac{(\delta+1)B}{\delta(\lambda+1)^2} - \frac{(\delta-\lambda)\ln\left(\frac{\lambda-k}{1-k\lambda}\right)}{\delta(\lambda+1)rB\left(\frac{1-k}{1+k}\right)} - \frac{B^2}{2(\lambda+1)^2}$$

$$- \frac{1}{2}r\sigma^2 B^2 - \frac{1}{2r}\ln\left(\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right). \tag{33}$$

As with gEPR, since $\rho = 1$, varying k has no effect on the diversification of the risk from the shocks to performance. Consequently, the only effects of varying k, holding aggregate effort fixed, are the negative effects stemming from the increase in the gap between efforts on the two tasks as k increases: A larger effort gap $\overline{e} - \underline{e}$ directly reduces the principal's benefit whenever $\delta > \lambda$ (the second term in (33)) and also results in the agent bearing more risk from the randomization (the final term in (33)). With $\rho = 1$, it is therefore optimal under gEAR, for any level of aggregate effort to be induced, to set k as small as possible, so as to induce as small a gap in efforts as possible. With k set arbitrarily close to, but larger than, -1, the agent is induced to choose a gap in efforts arbitrarily close to, but larger than, 0 (as shown by Proposition 3), and the principal achieves profits arbitrarily close to[12]

$$\Pi^{gEAR}(B) = \frac{\delta+1}{\delta}\frac{B}{(\lambda+1)^2} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2 B^2. \tag{34}$$

This is the same profit expression as arose under gEPR for any value of $k \in (-1, \frac{1}{\lambda})$. With B chosen optimally along with k, the principal can therefore, when $\rho = 1$, achieve profit under gEAR arbitrarily close to the maximized profit under gEPR, as given in equation (32).

Proposition 10 *Consider the limiting case of perfect correlation of the shocks: $\rho \to 1$. For any given level of aggregate effort, $\overline{e} + \lambda\underline{e}$, to be induced:*

1. *under gEPR, the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any λ and for any $k \in (-1, \frac{1}{\lambda})$, and any $k \in (-1, \frac{1}{\lambda})$ is optimal;*

2. *under gEAR, the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any λ as $k \to -1^+$, and the optimal value of $k \to -1^+$;*

There exists a critical δ above which gEAR and gEPR with k adjusted optimally are more profitable than the best deterministic scheme and below which the principal's most profitable incentive scheme is a symmetric deterministic menu (SDM).

This critical δ equals $\hat{\delta}(\lambda, R, \rho)$ defined in Proposition 9, evaluated at $\rho = 1$ and $R = r\sigma^2$.

---

[12] As k is lowered, the incentive coefficient β must be raised to keep aggregate effort, which is proportional to $B \equiv \beta(1+k)$, fixed.

5.4 Discussion

We have identified three environments in which random contracts, when designed optimally, can be shown to dominate the best deterministic scheme. In all three environments, gEAR and gEPR, with the parameter k adjusted optimally, both induce both types of agent to choose perfectly balanced efforts, and gEAR is as profitable for the principal as gEPR. In each setting, there is a critical degree of complementarity of the tasks for the principal above which the optimally designed randomized schemes dominate all deterministic schemes and below which the best scheme is the symmetric deterministic menu (SDM).

In none of these settings is the asymmetric deterministic menu (ADM) ever strictly the most profitable incentive scheme for the principal. There are two reasons for this finding. First, the ADM induces one type of agent to choose perfectly balanced efforts, while inducing the other type to choose fully focused efforts (and optimally insuring this type). Its outcome is therefore intermediate between the perfectly balanced efforts from both types induced by the randomized schemes and the fully focused efforts (with optimal insurance) generated by the SDM. Second, whenever $\lambda > 1$, the ADM cannot induce balanced efforts from one type without leaving informational rents to the other type. When $\lambda > 1$, these rents make the ADM strictly less profitable, for all values of δ, than the best alternative scheme, and only when $\lambda \to 1^+$ is the ADM as attractive as gEAR/gEPR and the SDM at the critical $\delta^S(r\sigma^2, \rho)$ where these other schemes are equally profitable.

The critical values of δ above which randomized schemes were optimal were closely related across the three environments, as Propositions 8, 9, 10 showed. Randomized schemes are more likely to be optimal the smaller is $r\sigma^2$: this reflects the imposition of greater risk on the agent by gEAR/gEPR than SDM per unit of aggregate effort induced. In the limit as $r\sigma^2 \to 0$, the critical values of δ approach λ: with risk no longer having a cost in this limit, the principal prefers randomized to deterministic schemes according to whether or not the perfectly balanced efforts induced by the randomized schemes are socially efficient. Randomized schemes are also more likely to be optimal the smaller is ρ, since a smaller ρ both i) reduces the gains from optimal insurance offered by the SDM and ii) reduces the risk costs under the randomized schemes of basing compensation on both performance measures. Finally, the critical values of δ are increasing in λ, since as λ increases, for any given level of aggregate effort to be induced, the gap between the cost of the risk imposed on the agent under randomized and deterministic schemes widens.

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

$$c(\overline{e}, \underline{e}) = \frac{1}{2}\left(\overline{e}^2 + 2s\lambda\overline{e}\underline{e} + \lambda^2\underline{e}^2\right) \tag{35}$$

where $\overline{e}$ 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 $\beta = \frac{\partial c}{\partial \overline{e}} = \frac{\partial c}{\partial \underline{e}}$, then

$$\overline{e}^{SD} = \frac{\beta(\lambda^2 - s\lambda)}{\lambda^2(1 - s^2)} \text{ and } \underline{e}^{SD} = \frac{\beta(1 - s\lambda)}{\lambda^2(1 - s^2)}. \tag{36}$$

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 $U(w - c(\overline{e}, \underline{e}))$, where $c(\overline{e}, \underline{e})$ has the generalized form in (35), 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

$$\beta = \frac{\partial c}{\partial \overline{e}} + \frac{\partial c}{\partial \underline{e}} \tag{37}$$

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

$$\frac{\frac{\partial c}{\partial \underline{e}}}{\frac{\partial c}{\partial \overline{e}}} = \frac{E\left[U'(\cdot)I_{\{\underline{x} \text{ is rewarded}\}}\right]}{E\left[U'(\cdot)I_{\{\overline{x} \text{ is rewarded}\}}\right]}, \tag{38}$$

32

where $\overline{x}$ 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 (38) is

$$\frac{\int U'(\alpha + \beta(\underline{e} + \underline{u}) - c(\overline{e}, \underline{e}))g(\underline{u})d\underline{u}}{\int U'(\alpha + \beta(\overline{e} + \overline{u}) - c(\overline{e}, \underline{e}))g(\overline{u})d\overline{u}}. \tag{39}$$

For ex post randomization, the right-hand side of (38) is

$$\frac{\int \int U'(\alpha + \beta(\underline{e} + \underline{u}) - c(\overline{e}, \underline{e}))I_{\{\overline{e}+\overline{u} \geq \underline{e}+\underline{u}\}}f(\overline{u}, \underline{u})d\overline{u}d\underline{u}}{\int \int U'(\alpha + \beta(\overline{e} + \overline{u}) - c(\overline{e}, \underline{e}))I_{\{\overline{e}+\overline{u} < \underline{e}+\underline{u}\}}f(\overline{u}, \underline{u})d\overline{u}d\underline{u}}. \tag{40}$$

Under a symmetric deterministic scheme, an interior optimal solution for efforts still satisfies $\beta = \frac{\partial c}{\partial \overline{e}} = \frac{\partial c}{\partial \underline{e}}$, so just as in the exponential-normal model, an interior solution is given by (36) and exists if and only if $s\lambda < 1$.

Proposition 11 *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*

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

and

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

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 $\frac{\overline{e}^{EAR}}{\underline{e}^{EAR}}$ and $\frac{\overline{e}^{EPR}}{\underline{e}^{EPR}}$ increase continuously from 1, whereas $\frac{\overline{e}^{SD}}{\underline{e}^{SD}}$ 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 $\overline{e} - \underline{e}$ 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 $\overline{e} - \underline{e}$.

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

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

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, (38) becomes

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

and for ex post randomization, (38) becomes

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

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

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

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.

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 optimally designed randomized contracts dominate all deterministic incentive schemes. We identified three such environments. Each of these environments has the feature that optimally designed randomized 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 the limiting case where the agents' risk aversion becomes infinitely large and the variance of the shocks to outputs becomes arbitrarily small. The final setting is that where the shocks affecting measured performance on the tasks become perfectly correlated. In all three of these environments, we showed that there is a critical degree of complementarity of tasks for the principal above which the optimal incentive scheme is a randomized one.

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 ad-

vance. The outcome under ex ante randomization is equivalent to the equilibrium outcome of a simultaneous-move game between the principal and the agent, and ex post randomization allows the principal to retain discretion over which performance measure to reward until after observing outputs. 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 with limited 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*, March 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. The comparison between SDM and ADMU involves comparing Π^{SDM} and Π^{ADMU}, where the former dominates if and only if

$$(\lambda + 1) \sqrt{\frac{\lambda^2 + R\left(\lambda^2 + 2\rho\lambda + 1\right)}{1 + R\left(1 - \rho^2\right)}} > 1 + \delta,$$

where the left-hand-side of the inequality is denoted $\delta^{SDM/ADMU} + 1$. The value $\delta^{ADMU/ADMC}$ solves

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

$\delta^{SDM/ADMU} \leq \delta^{ADMU/ADMC}$ if and only if

$$\lambda^2 \leq \frac{\lambda^2 + R\left(\lambda^2 + 2\rho\lambda + 1\right)}{1 + R\left(1 - \rho^2\right)},$$

and since $R \geq 0$, this is equivalent to

$$\lambda^2 \left(1 - \rho^2\right) \leq \lambda^2 + 2\rho\lambda + 1,$$

which is true for all $\rho \geq 0$.

We can also confirm that $\Pi^{ADMC} \geq \Pi^{SDM}$ for all $\delta > d^{ADMU/ADMC}$ since the critical value below which $\Pi^{SDM} > \Pi^{ADMC}$ is weakly less than $\delta^{ADMU/ADMC}$.

We thus have

$$\delta^{SDM/ADMU} - \lambda = (\lambda + 1)\left(\frac{\lambda^2 + R\left(\lambda^2 + 2\rho\lambda + 1\right)}{1 + R\left(1 - \rho^2\right)} - 1\right), \text{ and}$$

$$\delta^{ADMU/ADMC} - \lambda = (\lambda + 1)\left(\frac{\lambda^2 + R\left(\lambda^2 + 2\rho\lambda + 1\right)}{\lambda\left(1 + R\left(1 - \rho^2\right)\right)} - 1\right),$$

from which the rest of the proposition follows. ∎

Proof of Proposition 2. This proof demonstrates all the assertions in the statement of the proposition. In addition, it establishes that, if the principal could commit to arbitrary randomizing probabilities p and $1 - p$ such that with probability p, $\beta_1 = \beta$, $\beta_2 = 0$, and with probability $1 - p$, $\beta_1 = 0$ and $\beta_2 = \beta$, it would be optimal for the principal to commit to $p = \frac{1}{2}$, the value used throughout the text and in the statement of the proposition.

Agent 1 maximizes expected utility

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

The first-order conditions are

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

where

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

Adding the first-order conditions gives

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

$\Leftrightarrow$

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

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

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

For agent 2, analogous steps yield optimal effort levels

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

The maximized expected utilities are

$$
EU_1 = -p(\lambda+1)\exp\left(-r\left(\alpha+\frac{\beta^2}{2(\lambda+1)^2}-\frac{r\sigma^2\beta^2}{2}+\frac{\lambda\ln\frac{p\lambda}{1-p}}{r(\lambda+1)}\right)\right)
$$

$$
= -A_1\exp\left(-r\left(\alpha+\frac{\beta^2}{2(\lambda+1)^2}-\frac{r\sigma^2\beta^2}{2}\right)\right)
$$

$$
EU_2 = -(1-p)(\lambda+1)\exp\left(-r\left(\alpha+\frac{\beta^2}{2(\lambda+1)^2}-\frac{r\sigma^2\beta^2}{2}+\frac{\lambda\ln\frac{(1-p)\lambda}{p}}{r(\lambda+1)}\right)\right)
$$

$$
= -A_2\exp\left(-r\left(\alpha+\frac{\beta^2}{2(\lambda+1)^2}-\frac{r\sigma^2\beta^2}{2}\right)\right)
$$

where

$$
A_1 = p(\lambda+1)\exp\left(-\frac{\lambda\ln\frac{p\lambda}{1-p}}{\lambda+1}\right)
$$

$$
A_2 = (1-p)(\lambda+1)\exp\left(-\frac{\lambda\ln\frac{(1-p)\lambda}{p}}{\lambda+1}\right).
$$

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

$$
E[w] = \alpha+\frac{p}{2}\beta e_1^1+\frac{1-p}{2}\beta e_2^1+\frac{p}{2}\beta e_1^2+\frac{1-p}{2}\beta e_2^2
$$

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

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

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

$$
= \frac{\beta^2}{2(\lambda+1)^2}+\frac{r\sigma^2\beta^2}{2}+K(p,\lambda,r)
$$

where

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

The principal's expected profit is

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

and substituting for $E\left[w\right]$ yields

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

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 $K\left(p,\lambda,r\right)$ does not depend on β, the optimal choice of β is independent of the randomizing probability p. Furthermore, the principal's profit is increasing in p for $p \in \left(\frac{1}{1+\lambda},\frac{1}{2}\right)$, since

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

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

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

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

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

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

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

41

latter by $\underline{e}$, we have

$$\overline{e} + \lambda\underline{e} = \frac{\beta}{(\lambda+1)}$$

$$\overline{e} - \underline{e} = \frac{\ln\lambda}{r\beta}.$$

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 3. The proof follows along exactly the same lines as the proof of Proposition 2. ∎

Proof of Proposition 4. 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:

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

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

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

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

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

The first order condition with respect to e_1 is

$$
\begin{aligned}
0 &= -r(e_1 + \lambda e_2)m(-r\beta) \\
&\quad + r\beta \exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right) \Phi\left(\frac{e_2 - e_1 + r\beta\sigma^2(1-\rho)}{\theta}\right) \\
&\quad + \frac{1}{\theta} \exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right) \phi\left(\frac{e_2 - e_1 + r\beta\sigma^2(1-\rho)}{\theta}\right) \\
&\quad - \frac{1}{\theta} \exp\left(-r\beta e_2 + \frac{1}{2}r^2\beta^2\sigma^2\right) \phi\left(\frac{e_1 - e_2 + r\beta\sigma^2(1-\rho)}{\theta}\right).
\end{aligned} \tag{41}
$$

Similarly for e_2 we have

$$
\begin{aligned}
0 &= -\lambda r(e_1 + \lambda e_2)m(-r\beta) \\
&\quad + r\beta \exp\left(-r\beta e_2 + \frac{1}{2}r^2\beta^2\sigma^2\right) \Phi\left(\frac{e_1 - e_2 + r\beta\sigma^2(1-\rho)}{\theta}\right) \\
&\quad + \frac{1}{\theta} \exp\left(-r\beta e_2 + \frac{1}{2}r^2\beta^2\sigma^2\right) \phi\left(\frac{e_1 - e_2 + r\beta\sigma^2(1-\rho)}{\theta}\right) \\
&\quad - \frac{1}{\theta} \exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right) \phi\left(\frac{e_2 - e_1 + r\beta\sigma^2(1-\rho)}{\theta}\right).
\end{aligned} \tag{42}
$$

Adding the two first order conditions we find

$$
e_1 + \lambda e_2 = \frac{\beta}{(\lambda + 1)}. \tag{43}
$$

Expanding the third and fourth terms in the two first-order conditions (41) and (42) reveals that in both FOC's these terms net to 0 for all (e_1, e_2), and hence for (41) 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 (43) yields

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

$\Leftrightarrow$

$$
\lambda = \exp\left[r\beta(e_1 - e_2)\right] \frac{\Phi\left(\frac{e_1 - e_2 + r\beta\sigma^2(1-\rho)}{\theta}\right)}{\Phi\left(\frac{e_2 - e_1 + r\beta\sigma^2(1-\rho)}{\theta}\right)}. \tag{44}
$$

Both factors on the RHS of (44) 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 (44) 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 $\overline{e}^{EPR}$ and on his less-preferred task by $\underline{e}^{EPR}$. Define $d^{EPR} \equiv \overline{e}^{EPR} - \underline{e}^{EPR}$.

Using (43) and (44), we can express the maximized expected utility of both types under ex post randomization as

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

$\Leftrightarrow$

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

For both types of agent, the certainty equivalent is

$$ACE = \alpha + \beta\underline{e}^{EPR} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2\beta^2 - \frac{1}{r}\ln\left[\frac{1+\lambda}{\lambda}\Phi\left(\frac{\overline{e}^{EPR} - \underline{e}^{EPR} + r\beta\sigma^2(1-\rho)}{\theta}\right)\right] \quad (45)$$

while the principal's expected profit is

$$\begin{aligned}
\Pi^{EPR} &= \underline{e}^{EPR} + \frac{1}{\delta}\overline{e}^{EPR} - \alpha - \beta E\min\{x_1, x_2\} \\
&= \underline{e}^{EPR} + \frac{1}{\delta}\overline{e}^{EPR} - \alpha - \beta\left[\overline{e}^{EPR}\Phi\left(-\frac{d^{EPR}}{\theta}\right) + \underline{e}^{EPR}\Phi\left(\frac{d^{EPR}}{\theta}\right) - \theta\phi\left(\frac{d^{EPR}}{\theta}\right)\right] \\
&= \underline{e}^{EPR} + \frac{1}{\delta}\overline{e}^{EPR} - \alpha - \beta\underline{e}^{EPR} - \beta d^{EPR}\Phi\left(-\frac{d^{EPR}}{\theta}\right) + \beta\theta\phi\left(\frac{d^{EPR}}{\theta}\right). \quad (46)
\end{aligned}$$

Using (45) to substitute into (46) yields the principal's expected profit

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

∎

Proof of Proposition 5. The proof follows along exactly the same lines as the proof of Proposition 4. ∎

Proof of Proposition 6. Equations (14) and (23) give the principal's profit from interior effort choices by the agents under gEAR and gEPR, respectively, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$. The

44

proof proceeds in three steps:

Step 1:

$$\Pi^{gEPR}(\beta, k) \geq \underline{e}^{gEPR} + \frac{1}{\delta}\overline{e}^{gEPR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^g)^2\beta^2$$

This inequality reflects the fact that for any given β and k, gEPR imposes *lower* risk costs than would either of the deterministic contracts $w = \alpha + \beta x_1 + k\beta x_2$ or $w = \alpha + \beta x_2 + k\beta x_1$. To prove this inequality, we must show that the sum of the terms in the final three lines of equation (23) is non-negative. Define $d \equiv (1-k)(\overline{e} - \underline{e})$. Then the sum in question has the sign of

$$-r\beta d\Phi\left(\frac{-d}{\theta^g}\right) + r\beta\theta^g\phi\left(\frac{d}{\theta^g}\right) - \ln\left[\exp\{-r\beta d\}\Phi\left(\frac{-d}{\theta^g} + \frac{r\beta\theta^g}{2}\right) + \Phi\left(\frac{d}{\theta^g} + \frac{r\beta\theta^g}{2}\right)\right]. \tag{47}$$

Now define $t \equiv \frac{r\beta\theta^g}{2}$ and $y \equiv \frac{d}{\theta^g}$. The expression (47) can be rewritten as

$$h(y,t) \equiv -2ty\Phi(-y) + 2t\phi(-y) - \ln\left[\exp\{-2ty\}\Phi(-y+t) + \Phi(y+t)\right]. \tag{48}$$

It is not difficult to show, for all $y \geq 0$ and $t \geq 0$, that $h(y,t) \geq 0$ and, for future use, that $h(y,t)$ is decreasing in y.

Step 2: When $\delta \geq \lambda$,

$$\underline{e}^{gEPR} + \frac{1}{\delta}\overline{e}^{gEPR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^g)^2\beta^2$$

$$\geq \underline{e}^{gEAR} + \frac{1}{\delta}\overline{e}^{gEAR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^g)^2\beta^2$$

This step follows, when $\delta \geq \lambda$, from the facts that aggregate effort $\overline{e} + \lambda\underline{e}$ is equal under gEPR and gEAR and that the gap in efforts, $\overline{e} - \underline{e}$, is smaller under gEPR than gEAR.

Step 3:

$$\underline{e}^{gEAR} + \frac{1}{\delta}\overline{e}^{gEAR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^g)^2\beta^2$$

$$\geq \underline{e}^{gEAR} + \frac{1}{\delta}\overline{e}^{gEAR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^g)^2\beta^2$$

$$-\frac{1}{2r}\ln\left(\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right)$$

$$= \Pi^{gEAR}(\beta, k)$$

This step follows since, for any $\lambda \geq 1$ and $k \in (-1, \frac{1}{\lambda})$, $-\frac{1}{2r}\ln\left(\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right) \leq 0$. This reflects the fact that for any given β and k, gEAR imposes *higher* risk costs than would either of the deterministic contracts $w = \alpha + \beta x_1 + k\beta x_2$ or $w = \alpha + \beta x_2 + k\beta x_1$. ∎

Proof of Proposition 7. **Proof of Part 1:** For $\lambda = 1$, both gEAR and gEPR induce interior solutions for efforts for all $\beta > 0$ and $k \in (-1, 1)$. Therefore, from Proposition 6, we know that gEPR is more profitable than gEAR for any given (β, k), so it suffices to show that, for any given

(β, k), gEPR can be dominated in terms of profits by a suitably designed symmetric deterministic (SD) scheme.

For $\lambda = 1$, aggregate effort under gEPR is $\overline{e}^{gEPR} + \lambda \underline{e}^{gEPR} = \frac{\beta(1+k)}{2}$, and $\overline{e}^{gEPR} = \underline{e}^{gEPR} = \frac{\beta(1+k)}{4}$. Hence, for $\lambda = 1$, we can use equation (23) to write

$$\Pi^{gEPR}(\beta, k) = \frac{\delta + 1}{\delta}\frac{\beta(1+k)}{4} - \frac{1}{8}\beta^2(1+k)^2 - \frac{1}{2}r(\sigma^g)^2\beta^2 \tag{49}$$
$$- \frac{1}{r}\left(\ln\left[2\Phi\left(\frac{r\beta\theta^g}{2}\right)\right] - r\beta\theta^g\phi(0)\right).$$

Consider now a SD scheme with incentive coefficient β^{SD} chosen to induce the same level of aggregate effort as under gEPR for the given values of β and k:

$$\beta^{SD} = \frac{\beta(1+k)}{2}$$

Then, since $\lambda = 1$, $\overline{e}^{SD} = \underline{e}^{SD} = \frac{\beta(1+k)}{4}$, so SD also induces exactly the same effort levels on each task as gEPR. The principal's profit under the SD scheme is

$$\Pi^{SD}(\beta^{SD}) = \frac{\delta + 1}{\delta}\frac{\beta^{SD}}{2} - \frac{1}{2}\left(\beta^{SD}\right)^2 - r\sigma^2\left(\beta^{SD}\right)^2(1+\rho) \tag{50}$$
$$= \frac{\delta + 1}{\delta}\frac{\beta(1+k)}{4} - \frac{1}{8}\beta^2(1+k)^2 - \frac{1}{4}r\sigma^2\beta^2(1+k)^2(1+\rho)$$

Using equations (49) and (50) and the definitions of σ^g and θ^g in the statement of Proposition 5, we can write the difference in profits between the SD scheme and the gEPR scheme as

$$\Pi^{SD}(\beta^{SD}) - \Pi^{gEPR}(\beta, k) = \frac{r\beta^2}{2}\left[(\sigma^g)^2 - \frac{\sigma^2(1+k)^2)(1+\rho)}{2}\right] + \frac{1}{r}\left(\ln\left[2\Phi\left(\frac{r\beta\theta^g}{2}\right)\right] - r\beta\theta^g\phi(0)\right)$$
$$= \frac{1}{4}r\sigma^2\beta^2(1-\rho)(1-k)^2 + \frac{1}{r}\left(\ln\left[2\Phi\left(\frac{r\beta\theta^g}{2}\right)\right] - r\beta\theta^g\phi(0)\right).$$

Now, as in the proof of Proposition 6, define $t \equiv \frac{r\beta\theta^g}{2}$. Then

$$t^2 = \frac{r^2\sigma^2\beta^2(1-\rho)(1-k)^2}{2}, \tag{51}$$

and the profit difference given by equation (51) has the sign of

$$g(t) \equiv \frac{t^2}{2} + \ln[2\Phi(t)] - 2t\phi(0). \tag{52}$$

Analyzing this function we have

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

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,

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

Hence, since $\beta > 0$ and $\rho < 1$ imply that $t > 0$, we have shown that $\forall \beta > 0$ and $\rho < 1$, $\Pi^{SD}(\beta^{SD}) - \Pi^{gEPR}(\beta, k) > 0$ If $\rho = 1$, then $\theta^g = 0$ for all $k \in (-1, 1)$, so $t = 0$, hence $\Pi^{SD}(\beta^{SD}) - \Pi^{gEPR}(\beta, k) > 0$.

Proof of Part 2: We first show that if gEAR induces a corner solution for efforts for given (β, k), then it can be dominated in terms of profits by a suitably designed SD scheme. When gEAR induces a corner solution for efforts (so $\underline{e}^{gEAR} = 0$), $\bar{e}^{gEAR}$ satisfies the FOC

$$
\frac{\beta - \bar{e}^{gEAR}}{\bar{e}^{gEAR} - k\beta} = \exp\left\{r\beta\bar{e}^{gEAR}(1 - k)\right\}. \tag{53}
$$

Since the RHS of (53) is > 1 for $k < 1$, (53) implies that

$$
\bar{e}^{gEAR} < \frac{\beta(1 + k)}{2}. \tag{54}
$$

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

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

where $Z \equiv \frac{\beta - \bar{e}^{gEAR}}{\bar{e}^{gEAR} - k\beta} > 1$.

Consider now a SD scheme with incentive coefficient β^{SD} chosen to induce the same effort pair $\left(\bar{e}^{gEAR}, 0\right)$ as under gEAR for the given values of β and k: $\beta^{SD} = \bar{e}^{gEAR}$. The principal's profit under this SD scheme is

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

Therefore

$$\begin{aligned}
\Pi^{SD}\left(\beta^{SD}\right) - \Pi^{gEAR}\left(\beta, k\right) \;&>\; \frac{r\sigma^2}{2}\left[\beta^2(1+2\rho k + k^2) - 2\left(1+\rho\right)\left(\bar{e}^{gEAR}\right)^2\right] \\
&>\; \frac{\left(\bar{e}^{gEAR}\right)^2 r\sigma^2}{2}\left[\frac{4(1+2\rho k + k^2)}{(1+k)^2} - 2\left(1+\rho\right)\right] \\
&=\; \frac{\left(\bar{e}^{gEAR}\right)^2 r\sigma^2}{(1+k)^2}\left[(1-\rho)(1-k)^2\right] \\
&\geq\; 0,
\end{aligned}$$

where the first inequality holds since $Z > 1$ and the second follows from inequality (54).

We now show that if gEPR induces a corner solution for efforts for given (β, k), then it can be dominated in terms of profits by a suitably designed SD scheme. When gEPR induces a corner solution for efforts (so $\underline{e}^{gEPR} = 0$), $\bar{e}^{gEPR}$ satisfies the FOC

$$\frac{\beta - \bar{e}^{gEPR}}{\bar{e}^{gEPR} - k\beta} = \exp\left\{r\beta(1-k)\bar{e}^{gEPR}\right\}\frac{\Phi\left(\frac{(1-k)\bar{e}^{gEPR} + r(\sigma^g)^2\beta(1-\rho^g)}{\theta^g}\right)}{\Phi\left(\frac{-(1-k)\bar{e}^{gEPR} + r(\sigma^g)^2\beta(1-\rho^g)}{\theta^g}\right)}. \tag{55}$$

Since the RHS of (55) is ≥ 1 for $k < 1$, (55) implies that $\bar{e}^{gEPR} \leq \frac{\beta(1+k)}{2}$. When gEPR induces A to choose the corner solution $\left(\bar{e}^{gEPR}, 0\right)$,

$$\begin{aligned}
\Pi^{gEPR}(\beta, k) = \frac{\bar{e}^{gEPR}}{\delta} \;&-\; \frac{1}{2}\left(\bar{e}^{gEPR}\right)^2 - \frac{1}{2}r\beta^2(\sigma^g)^2 \\
&-\; \frac{1}{r}\ln\left[\Phi(+) + \exp\left\{-r\beta(1-k)\bar{e}^{gEPR}\right\}\Phi(-)\right], \\
&-\; \beta(1-k)\bar{e}^{gEPR}\Phi\left(\frac{-(1-k)\bar{e}^{gEPR}}{\theta^g}\right) + \beta\theta^g\phi\left(\frac{(1-k)\bar{e}^{gEPR}}{\theta^g}\right)
\end{aligned}$$

where

$$\begin{aligned}
\Phi(+) \;&\equiv\; \Phi\left(\frac{(1-k)\bar{e}^{gEPR} + r(\sigma^g)^2\beta\left(1-\rho^g\right)}{\theta^g}\right) \\
\Phi(-) \;&\equiv\; \Phi\left(\frac{-(1-k)\bar{e}^{gEPR} + r(\sigma^g)^2\beta\left(1-\rho^g\right)}{\theta^g}\right).
\end{aligned}$$

Consider now a SD scheme with incentive coefficient β^{SD} chosen to induce the same effort pair $\left(\bar{e}^{gEPR}, 0\right)$ as under gEPR for the given values of β and k: $\beta^{SD} = \bar{e}^{gEPR}$. The principal's profit under this SD scheme is

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

Therefore, $\Pi^{SD}\left(\beta^{SD}\right) - \Pi^{gEPR}\left(\beta, k\right)$ has the sign of

$$\frac{r^2\sigma^2}{4} \quad (2\beta^2(1 + 2\rho k + k^2) - 4(1 + \rho)\left(\bar{e}^{EPR}\right)^2) \tag{56}$$

$$+ \quad \ln\left[\Phi(+) + \exp\left\{-r\beta(1 - k)\bar{e}^{gEPR}\right\}\Phi(-)\right] \tag{57}$$

$$+ \quad r\beta(1 - k)\bar{e}^{gEPR}\Phi\left(\frac{-(1 - k)\bar{e}^{gEPR}}{\theta^g}\right) + r\beta\theta^g\phi\left(\frac{(1 - k)\bar{e}^{gEPR}}{\theta^g}\right). \tag{58}$$

Now define $y \equiv \frac{(1-k)\bar{e}^{gEPR}}{\theta^g}$ and $t \equiv \frac{r\beta\theta^g}{2}$. Then (58) can be rewritten as

$$\frac{t^2}{2} - h(y, t),$$

where, as in the proof of Proposition 6,

$$h(y, t) \equiv -2ty\Phi(-y) + 2t\phi(-y) - \ln\left[\exp\{-2ty\}\Phi(-y + t) + \Phi(y + t)\right].$$

In the proof of Proposition 6 it was noted that for all $y \geq 0$, $t \geq 0$, $h(y, t)$ is decreasing in y, and hence

$$\frac{t^2}{2} - h(y, t) \quad \geq \quad \frac{t^2}{2} - h(0, t)$$

$$= \quad \frac{t^2}{2} + \ln\left[2\Phi(t)\right] - 2t\phi(0)$$

$$= \quad g(t),$$

where the function $g(t)$ was defined in the proof of Part 1 of this proposition and was there shown to be strictly positive for all $t > 0$.

Proof of Part 3: There are two cases to consider: (i) gEAR and gEPR induce interior solutions for efforts or (ii) gEAR and gEPR induce corner solutions. The proof of Part 2 has dealt with the latter case, so here we treat the former. From Proposition 6, we know that gEPR is more profitable than gEAR for any given (β, k) when both schemes induce interior solutions for efforts, so it suffices to show that, when $\delta \leq \lambda$, for any given (β, k), gEPR can be dominated in terms of profits by a suitably designed symmetric deterministic (SD) scheme.

Define $d \equiv (1 - k)(\bar{e}^{gEPR} - \underline{e}^{gEPR})$, $y \equiv \frac{d}{\theta^g}$, and $t \equiv \frac{r\beta\theta^g}{2}$. Then from the proof of Proposition 6, we know that we can write

$$\Pi^{gEPR}(\beta, k) = \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}\left(\bar{e} + \lambda\underline{e}\right)^2 - \frac{1}{2}r\beta^2(\sigma^g)^2 + \frac{1}{r}h(y, t)$$

$$\leq \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}\left(\bar{e} + \lambda\underline{e}\right)^2 - \frac{1}{2}r\beta^2(\sigma^g)^2 + \frac{1}{r}h(0, t)$$

$$= \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}\left(\bar{e} + \lambda\underline{e}\right)^2 - \frac{1}{2}r\beta^2(\sigma^g)^2 + \frac{1}{2}[-\ln(2\Phi(t)) + 2t\phi(0)]$$

$$\leq \frac{\bar{e} + \lambda\underline{e}}{\delta} - \frac{1}{2}\left(\bar{e} + \lambda\underline{e}\right)^2 - \frac{1}{2}r\beta^2(\sigma^g)^2 + \frac{1}{r}\left[-\ln\left(2\Phi(t)\right) + 2t\phi(0)\right]$$

$$= \frac{\beta(1 + k)}{\delta(1 + \lambda)} - \frac{1}{2}\frac{\beta^2}{(1 + \lambda)^2} - \frac{1}{2}r\beta^2(\sigma^g)^2 + \frac{1}{r}\left[-\ln\left(2\Phi(t)\right) + 2t\phi(0)\right], \tag{59}$$

where the first inequality follows from the fact that $h(y, t)$ is decreasing in y and the second from

the fact that, by assumption, $\delta \leq \lambda$.

Consider now a SD scheme with incentive coefficient β^{SD} chosen to induce the same aggregate effort as under gEPR for the given values of β and k: $\beta^{SD} = \frac{\beta(1+k)}{1+\lambda}$. The principal's profit under this SD scheme is

$$\Pi^{SD}(\beta^{SD}) = \frac{\beta(1+k)}{\delta\,(1+\lambda)} - \frac{1}{2}\frac{\beta^2(1+k)^2}{(1+\lambda)^2} - r\sigma^2\beta^2(1+k)^2\frac{1+\rho}{(1+\lambda)^2} \tag{60}$$

Hence from (59) and (60) we can conclude that

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

where the first inequality is a consequence of the inequalities in (59), the second inequality follows from the fact that $\lambda \geq 1$, the second equality uses (51), and the final line uses the definition of $g(t)$ in (52) and its non-nonegativity, as was established in the proof of Part 1 of this proposition. $\blacksquare$