Gaming and Strategic Ambiguity in Incentive Provision*

Florian Ederer† Richard Holden‡ Margaret Meyer§
UCLA Chicago and NBER Oxford and CEPR

December 15, 2011

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 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 ambiguous schemes impose more risk on the agent per unit of effort induced. By identifying settings in which ambiguous schemes are especially effective in inducing balanced efforts, we show that, if tasks are sufficiently complementary for the principal, ambiguous incentive schemes can dominate the best deterministic scheme and can completely eliminate the efficiency losses from the agent's better knowledge of the environment. (*JEL* L13, L22)

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

*We thank Philippe Aghion, Simon Board, Patrick Bolton, Jeremy Bulow, Kate Doornik, Guido Friebel, Robert Gibbons, Edward Glaeser, Oliver Hart, Bengt Holmström, Ian Jewitt, Navin Kartik, Shmuel Leshem, Steven Lippman, Steven Matthews, John Moore, In-Uck Park, Charles Roddie, Julia Simon-Kerr, Jesse Shapiro, Jason Snyder, Alexander Stremitzer, Jeroen Swinkels, and Jean Tirole for helpful discussions and suggestions, as well as seminar participants at Bristol, Chicago, Columbia, Frankfurt, Gerzensee, LSE, MIT, Oxford, Rotterdam, UCL, UCLA, and USC Law School.

†UCLA Anderson School of Management, 110 Westwood Plaza, Cornell Hall, D515, Los Angeles, CA 90095-1481, U.S.A., florian.ederer@anderson.ucla.edu.

‡University of Chicago Booth School of Business, 5807 S. Woodlawn Ave., Chicago, IL, 60637, U.S.A., richard.holden@chicagobooth.edu.

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

is also an important practical matter. There is a large informal literature which documents the perverse effects of (high-powered) incentive schemes. Much of this literature concludes that unless the incentive designer can measure performance on 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 light of apparent gaming of incentive schemes that were designed to reduce patient waiting times (Bevan and Hood 2004).[2] 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. Gaming is also pervasive in law school rankings as reported prominently in the popular press (e.g., "Law School Rankings Reviewed to Deter 'Gaming'", Wall Street Journal, August 26, 2008). In fact, "it's become an open secret that many law-school deans strategize specifically to improve their rank in U.S. News, to try to reap more interest by employers in their students, energize alumni donors and to put their own job out of jeopardy".[3] In response, U.S. News considered changing its widely read law school ranking to deter such gaming. Similarly, law scholars (e.g., Osler 2010) have argued against transparency in law school rankings methodology to prevent schools from gaming the same rankings that "legal academics almost uniformly criticize". There are numerous other examples in different, but related, incentive provision problems. The locations of speed cameras are often randomized,[4] 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

[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] Dranove, Kessler, McClellan and Satterthwaite (2003) also document that in the US report cards for hospitals encourage providers to "game" the system by avoiding sick patients or seeking healthy patients.

[3] Standard measures include cutting the number of full-time students to boost the median LSAT scores and GPAs and adding more part-time students, make-work jobs for own graduates to bump up employment numbers and spending more heavily to promote faculty scholarship to U.S. News voters than supporting the production of that scholarship. These measures have significant effects on rankings. For example, the University of Baltimore Law School rose from #170 to #125 in one year and its dean admitted that "U.S. News is not a moral code, it's a set of seriously flawed rules of a magazine, and I follow the rules without hiding anything".

[4] See Lazear (2006) for a model of this and related phenomena.

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 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.[5] Imposing less risk on the agent allows the principal to provide higher-powered incentives. In our model, however, ambiguity, 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; 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, even when menus of contracts are used as screening devices, because at least one type of agent would always want to choose the "wrong" contract. In this setting, it is advantageous to consider a richer contracting space, including ambiguous contracts. Such contracts make compensation ambiguous from the point of view of the agent, in that he knows that the compensation schedule ultimately used will take one of two possible forms, but at the time he chooses his efforts, he does not know which form will be used. The two possible compensation schedules differ with respect to which performance measure is more highly rewarded. Under *ex ante randomization*, the principal chooses randomly, before outputs are observed, which compensation schedule to employ. Under *ex post discretion*, the principal chooses which compensation schedule to employ 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 compensation schedule. Under ex post discretion, 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. For both types of ambiguous incentive scheme, we show how the principal can use the relative weight on the two performance

[5] 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.

measures to adjust the intensity of the incentives provided to the agent to choose balanced efforts.

Our analysis shows that ambiguous contracts are more robust to uncertainty about the agent's preferences than are deterministic ones. Specifically, we show that with ambiguous 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.

The benefits of ambiguous incentive schemes in deterring gaming do, nevertheless, come at a cost: they impose more risk on the agent. 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 ambiguous scheme. In general, therefore, the principal faces a trade-off between the stronger incentives for balanced efforts that arise under ambiguous schemes and the better insurance that is provided by deterministic schemes.

Our key contribution is to identify settings in which optimally designed ambiguous contracts dominate all deterministic incentive schemes. We identify three such environments. Each of these environments has the feature that optimally designed ambiguous 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 optimally designed ambiguous schemes allow the principal to achieve as high a payoff as if he knew the agent's preferences across tasks. That is, in these environments, ambiguous incentive schemes completely eliminate the efficiency losses from the agent's hidden information.

1.1 Related Literature

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.[6] 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

[6] A different strand of literature documents perverse incentives attributed to nonlinear schemes whereby agents make intertemporal effort shifts (e.g., Asch (1990), Oyer (1998) and Larkin (2006) among many others).

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

Bernheim and Whinston (1998) analyze the incompleteness of observed contracts, a phenomenon they term "strategic ambiguity" and 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 focus on explaining optimal contractual incompleteness, rather than incentive provision in a moral hazard setting. Jehiel (2011) shows in an abstract moral hazard setup that it can be desirable that some aspects of the interaction between principal and agent (e.g., incentive schemes governing their team coworkers and how important tasks are) be kept unknown to the workers. Finally, Jehiel and Newman (2009) study the evolution of incentive systems. Principals may offer either loophole contracts that deter only the harmful actions they deem sufficiently likely to exist (and agents then cheat) or loophole-free contracts that deter all cheating, thereby conveying little information about feasible actions to other principals. The result is cycling of contract types that alternately deter and encourage undesired behavior.

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 ambiguous 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.[7] The benefit of randomization derives

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

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 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, ambiguous contacts *cannot* dominate the best deterministic contract. Thus, the attractiveness of ambiguous 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 ambiguous contracts optimal. Second, the ambiguous contracts in our model do not require the principal to commit to the randomizing procedure in 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 discretion allows the principal to retain discretion over which performance measure to reward until after observing outputs. Therefore, our ambiguous schemes are feasible even when the principal is unable to commit to complicated non-linear contracts. *Indeed, we speculate that much of the appeal of ambiguous contracts is that they approximate the outcomes of 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 analyze the form of the optimal deterministic incentive scheme and the payoff it generates for the principal. Section 4 studies ex ante randomization and ex post discretion. Section 5 identifies settings in which ambiguous contracts are dominated by one or more of the deterministic schemes. Section 6, which is the heart of the paper, identifies environments in which ambiguous contracts dominate the best deterministic scheme. Section 7 shows that our results are robust to various extensions such as relaxing the assumption that tasks are perfect substitutes for the agent. Proofs not provided in the text are contained in the appendix.

2 The Model

A principal hires an agent to perform a job for her. The agent's performance on the job has two distinct dimensions, which we term "tasks". 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

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

where 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[8]: $\frac{\partial c_i / \partial e_1}{\partial c_i / \partial e_2}$ does not vary with (e_1, e_2). Nevertheless, since $\lambda \geq 1$, each type of agent is biased towards a preferred task: for the type-i agent, the marginal cost of effort on task i is (weakly) lower than the marginal cost of effort on the other task. We assume that both types of agent have an exponential von Neumann-Morgenstern utility function with coefficient of absolute risk aversion r, so the type-i agent's utility function is

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

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

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

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

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.[9]

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.

Below we consider a variety of incentive schemes. Throughout the analysis, we restrict attention to compensation schedules in which, ex post, after all choices are made and random variables are

[8] In Section 5, we show that our key results are robust to a relaxation of this assumption.

[9] We assume throughout that difficulties of coordination would make it prohibitively costly for the principal to split the job between two agents and induce each agent to focus on a single dimension (task).

realized, the agent's payment is a linear and separable function of the performance measures:

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

We will say an incentive scheme (possibly involving menus) is deterministic if, at the time the agent signs the contract or makes his choice from the menu, he is certain about what values of α, β_1, and β_2 will be employed in determining his pay. If, instead, even after making his choice from a menu, the agent is uncertain about the values of α, β_1, or β_2, we will say that the incentive scheme is ambiguous.

In the next section, we study deterministic incentive schemes and show 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 $r\sigma^2$ (which represents the importance of risk aversion under deterministic schemes).[10]

Section 4 introduces the two classes of ambiguous schemes on which we focus. A contract with *ex ante randomization* (EAR) 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 $k \in (-1, 1)$. Under this scheme, the principal commits to employ a randomizing device to determine whether the agent's pay will be more sensitive to performance on task 1 or task 2. Thus the agent is 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. Under a contract with *ex post discretion* (EPD), the principal, *after* observing the performance measures x_1 and x_2, chooses whether to pay the agent $w = \alpha+\beta x_1+k\beta x_2$ or $w = \alpha+\beta x_2+k\beta x_1$, where again $k \in (-1, 1)$. Under both classes of ambiguous 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 discretion do the agent's efforts influence which set of weights is ultimately used.

3 Deterministic Contracts

3.1 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 linear contract can take one of two possible forms. The first form is a symmetric contract, with $\beta_1 = \beta_2 = \beta$: $w = \alpha + \beta x_1 + \beta x_2$. 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

[10] For deterministic schemes, 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 ambiguous schemes, r and σ^2 have separate influences on the agent's effort choices and therefore on the principal's payoff.

form of contract rewards performance on only one task (which we take to be task 1 below) and so induces the agent to exert effort on only one task, and it 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 non-negative effort pairs that equate the common marginal cost of effort on the two tasks, $e_1 + e_2$, to the common marginal benefit, β. Since the parameter δ in the principal's benefit function is greater than or equal to one, the principal prefers the agent to choose $e_1 = e_2 = \frac{\beta}{2}$, and we assume that the agent does indeed choose this balanced effort allocation.

Lemma 1 *There exists a threshold $\delta^1(r\sigma^2, \rho) > 1$ which is increasing in each of its arguments, such that for δ above $\delta^1(.)$, an optimally designed SD contract is optimal and for δ below $\delta^1(.)$, an optimally designed OT contract is optimal.*

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, $r\sigma^2\beta^2(1+\rho)$, is larger than that under the OT contract, $\frac{1}{2}r\sigma^2\beta^2\left(1-\rho^2\right)$. Therefore the principal faces a tradeoff between the more balanced efforts induced by SD and the lower risk imposed by OT. Lemma 1 shows that there is a critical value $\delta^1(r\sigma^2, \rho)$ of the principal's complementarity parameter δ above which the SD contract is optimal and below which the OT contract is preferred.

3.2 Two Types of Agent but No Hidden Information

As a benchmark for the subsequent analysis, suppose now that $\lambda > 1$, so there are two distinct types of agent, but that the principal can observe the agent's type and offer each type a different contract. We will refer to this setting as the "no hidden information benchmark" (NHI).

In this setting, the optimal pair of contracts (one for each type of agent) can take one of two possible forms (analogous to the two possible optimal contracts for the case when $\lambda = 1$). The first form, which we denote (C_1^{bal}, C_2^{bal}), induces each type of agent to choose perfectly balanced efforts:

$$\begin{aligned} C_1^{bal} &: \quad w_1 = \alpha + \beta x_1 + \lambda \beta x_2, \\ C_2^{bal} &: \quad w_2 = \alpha + \beta x_2 + \lambda \beta x_1. \end{aligned}$$

Since C_i^{bal}, the contract assigned by the principal to the agent of type i, rewards agent i's more costly task at a rate λ times as large as his less costly task, agent i is indifferent over all non-negative effort pairs satisfying $\beta = e_i + \lambda e_j$, so is willing to choose $e_i = e_j = \frac{\beta}{1+\lambda}$. As λ approaches 1, for both types of agent the contract C_i^{bal} approaches the symmetric deterministic (SD) contract studied in the previous subsection.

The second type of contract pair which can be optimal in the "no hidden information benchmark" setting is a mirror-image pair of "one task" (OT) contracts, which induce each type of agent to exert effort only on his preferred task and use performance on the other task only to improve insurance:

$$\begin{aligned} C_1^{OT} &: \ w_1 = \alpha + \beta x_1 - \rho\beta x_2, \\ C_2^{OT} &: \ w_2 = \alpha + \beta x_2 - \rho\beta x_1. \end{aligned}$$

When assigned contract C_i^{OT}, each agent i chooses $e_i = \beta$ and $e_j = 0$. The principal's payoff will be the same from each type of agent, and with α and β chosen optimally. Note that this payoff from the contract pair (C_1^{OT}, C_2^{OT}) is independent of λ, since neither type of agent exerts any effort on his more costly task.

In choosing between the contract pairs (C_1^{bal}, C_2^{bal}) and (C_1^{OT}, C_2^{OT}) in the NHI setting, the principal faces a tradeoff between the more balanced efforts induced by the former pair and the lower risk cost imposed by the latter pair. [**FLORIAN: Do we want a figure like in the slides here? I don't think we do.]

Lemma 2 *There exists a threshold*

$$\delta^{NHI}(\lambda, r\sigma^2, \rho) \equiv (\lambda + 1)\left[\frac{1 + r\sigma^2(1 + 2\rho\lambda + \lambda^2)}{1 + r\sigma^2(1 - \rho^2)}\right]^{\frac{1}{2}} - 1 \tag{1}$$

such that for δ above $\delta^{NHI}(.)$, an optimally designed balanced (C_1^{bal}, C_2^{bal}) contract is optimal and for δ below $\delta^{NHI}(.)$, an optimally designed OT (C_1^{OT}, C_2^{OT}) contract is optimal. $\delta^{NHI}(.)$ is increasing in each of its arguments and, as $\lambda \to 1$, $\delta^{NHI} \to \delta^1$.

3.3 The General Case: Two Types of Privately-Informed Agent

In the general case where $\lambda > 1$ and the agent is privately informed about his preferences across tasks, the principal can use menus of contracts as screening devices. In principle, three possible patterns of effort could emerge: (a) both types of agent could exert balanced efforts, (b) one type could exert balanced efforts and the other type focused effort, or (c) both types could exert focused effort. We now analyze the form of optimally designed menus of contracts, the patterns of effort they induce, and the payoffs they generate for the principal.

Lemma 3 *When $\lambda > 1$, no menu of linear deterministic contracts can induce both types of agent to choose strictly positive efforts on both tasks.*

Lemma 3 is straightforward. If such a menu existed, it would have the form

$$\begin{aligned} C_1 &: \ w_1 = \alpha_1 + \beta_1 x_1 + \lambda\beta_1 x_2, \\ C_2 &: \ w_2 = \alpha_2 + \beta_2 x_2 + \lambda\beta_2 x_1. \end{aligned}$$

But at least one type of agent would select the "wrong" contract from the menu and exert effort only on his preferred task. Therefore, we can confine attention to a contract that is either (b) an "asymmetric deterministic menu" (ADM)

$$\begin{aligned} C_1^{ADM} &: \ w_1 = \alpha_1 + \beta_1 x_1 - \rho\beta_1 x_2, \\ C_2^{ADM} &: \ w_2 = \alpha_2 + \beta_2 x_2 + \lambda\beta_2 x_1. \end{aligned}$$

or (c) a "symmetric deterministic menu" (SDM)

$$\begin{aligned} C_1^{SDM} &: \quad w_1 = \alpha + \beta x_1 - \rho\beta x_2, \\ C_2^{SDM} &: \quad w_2 = \alpha + \beta x_2 - \rho\beta x_1. \end{aligned}$$

Under the ADM, one agent (here, type-2 agent) chooses balanced efforts while the coefficient on x_2 in C_1^{ADM} exploits the correlation in the shocks to the performance measure to improve the insurance offered to agent 1. Under the SDM, both agents exert fully-focused efforts and the coefficient for the output on their less preferred task is used for insurance. Relative to the SDM, the ADM has the benefit that it induces one type of agent to choose balanced efforts, but it has the costs that it imposes more risk on that agent type and also necessitates leaving rents to the other agent type. Whether the benefit of the ADM outweighs its costs depends on δ, the degree of complementarity between the tasks in the principal's benefit function.

Proposition 1 *(i) When $\delta > \delta^{NHI}$, the principal is strictly better off when hidden information is absent than when it is present.*
(ii) When the agent is privately informed about his preferences, there exists a critical threshold

$$\delta^{HI}(\lambda, r\sigma^2, \rho) \equiv (\lambda + 1)\left[\frac{\lambda^2 + r\sigma^2(1 + 2\rho\lambda + \lambda^2)}{1 + r\sigma^2(1 - \rho^2)}\right]^{\frac{1}{2}} - 1, \tag{2}$$

such that the principal's payoff is maximized by an optimally designed ADM if $\delta > \delta^{HI}$, whereas for $\delta \leq \delta^{HI}$ the principal's payoff is maximized by an optimally designed SDM. $\delta^{HI}(\lambda, r\sigma^2, \rho)$ is increasing in each of its arguments, $\delta^{HI} > \delta^{NHI}$ for all $\lambda > 1$ and, as $\lambda \to 1$, $\delta^{HI} \to \delta^1$.

Let us now compare the results from the analyses of the settings with and without hidden information on the agent's part. First, it is clear that, whenever $\delta > \delta^{NHI}$, the principal is strictly better off when hidden information is absent than when it is present: when $\delta > \delta^{NHI}$, the principal in the "no hidden information benchmark" opts to induce both types of agent to choose perfectly balanced efforts, an outcome which is infeasible under hidden information. Second, comparing (2) with (1) shows that for all $\lambda > 1$, $\delta^{HI}(\lambda, r\sigma^2, \rho) > \delta^{NHI}(\lambda, r\sigma^2, \rho)$, while as λ approaches 1, these two critical values of δ become equal, at $\delta^1(r\sigma^2, \rho)$. In the presence of hidden information on the agent's part, the principal's complementarity parameter δ must be larger than when hidden information is absent in order for her to find it optimal to attempt to induce balanced efforts (even from just one type of agent). The reason is the rents that hidden information forces the principal to leave to one agent type when he uses the ADM to induce balanced efforts from the other type. Finally, under hidden information, for values of $\delta > \delta^1(r\sigma^2, \rho)$, the principal's maximized payoff drops discontinuously as λ is increased from 1 (where a SD contract is optimal) to a value slightly greater than 1, where the optimal scheme is an ADM. This discontinuous drop reflects the impossibility, for even a small degree of preference across tasks, of inducing balanced efforts from both agent types with a deterministic incentive scheme. In contrast, in the "no hidden information benchmark", where it is feasible to induce balanced efforts from both agent types for all values of λ, the principal's maximized payoff falls continuously as λ is increased from 1. Figure 1 summarizes the conclusions from our analysis of the optimal deterministic schemes.

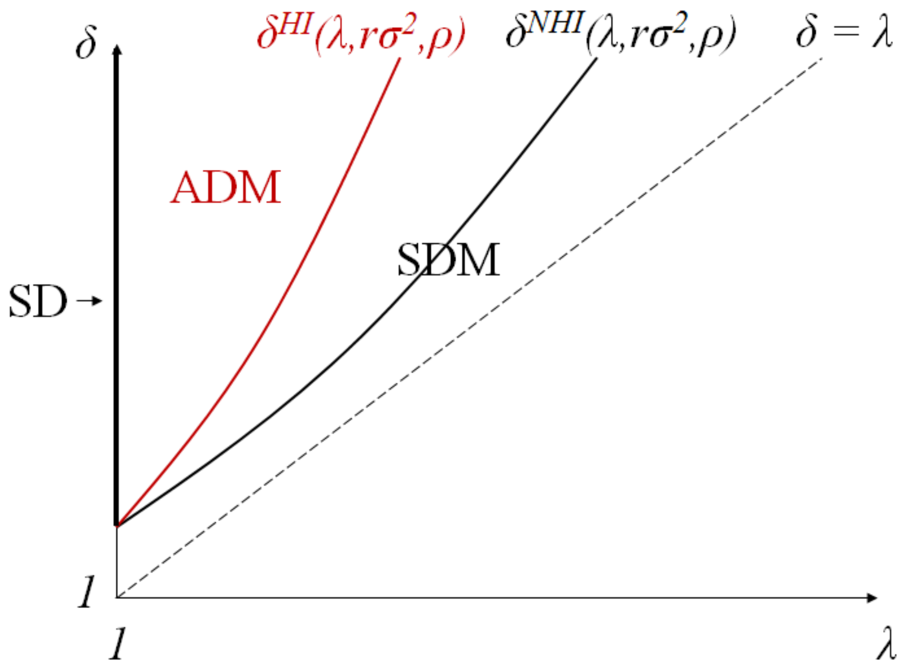

Figure 1: Optimal deterministic schemes under hidden information

4 Ambiguous Contracts

4.1 Ex Ante Randomization

A contract with *ex ante randomization* (EAR) 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 $k \in (-1, 1)$. Under such a scheme, the principal commits to employ a randomizing device to determine whether the agent's pay will be more sensitive to performance on task 1 or task 2. If the agent chooses unequal efforts on the two tasks, the principal's use of randomization exposes the agent to extra risk, risk against which he can insure himself by choosing more balanced efforts. By varying the level of k, the principal can affect how much risk the randomization per se imposes on the agent and thereby affect the strength of the agent's incentives to balance his efforts. 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 which, whenever $\lambda > 1$, induces both types of agent to exert effort only on their preferred task. The smaller is k, the greater is the risk imposed on the agent by the principal's randomization, so intuitively the stronger are the agent's incentives to self-insure by choosing a more balanced profile of efforts.

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

$$\overline{e}^{EAR} + \lambda \underline{e}^{EAR} = \frac{\beta(1+k)}{\lambda+1} \tag{3}$$

$$\overline{e}^{EAR} - \underline{e}^{EAR} = \frac{\ln \frac{\lambda - k}{1 - k\lambda}}{r\beta(1-k)}. \tag{4}$$

(iii) The gap in efforts, $\overline{e}^{EAR} - \underline{e}^{EAR}$, 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 $k \to -1^+$.
(iv) The principal's payoff from interior effort choices by the agents under EAR, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$, is

$$\begin{aligned} \Pi^{EAR}(\beta, k) = \underline{e}^{EAR} + \frac{1}{\delta}\overline{e}^{EAR} \quad &- \quad \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} \\ &- \quad \frac{1}{2} r\sigma^2\beta^2(1 + 2\rho k + k^2) - \frac{1}{2r}\ln\left(\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right) \end{aligned} \tag{5}$$

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.

To understand equation (3), note first that the sum of the expected marginal monetary returns to effort on the two tasks must be $\beta(1+k)$, since it is certain that one task will be rewarded at rate β and the other at rate $k\beta$. If optimal efforts for the agents are interior, then adding the first-order conditions for effort on the two tasks must yield $\beta(1+k) = \partial c/\partial \overline{e} + \partial c/\partial \underline{e}$ for both types of agent. Since $\partial c/\partial \overline{e} + \partial c/\partial \underline{e} = (1+\lambda)(\overline{e} + \lambda \underline{e})$, this gives us equation (3). Throughout the paper, we will refer to the quantity $\overline{e} + \lambda \underline{e}$ as an agent's *aggregate effort*, since it is the quantity which determines his total cost of effort.

Equation (4) is derived by substituting (3) into either of the first-order conditions for optimal efforts, which yields

$$\frac{\lambda - k}{1 - k\lambda} = \frac{E\left[U'(\cdot) I_{\{\underline{x} \text{ is rewarded more highly than } \overline{x}\}}\right]}{E\left[U'(\cdot) I_{\{\overline{x} \text{ is rewarded more highly than } \underline{x}\}}\right]}, \tag{6}$$

where $\overline{x}$ (respectively, $\underline{x}$) denotes performance on an agent's less costly (respectively, more costly) task. Since under EAR, each of the two possible compensation schemes is employed with probability one-half,

$$\frac{E\left[U'(\cdot) I_{\{\underline{x} \text{ is rewarded more highly than } \overline{x}\}}\right]}{E\left[U'(\cdot) I_{\{\overline{x} \text{ is rewarded more highly than } \underline{x}\}}\right]} = \exp\left[r\beta(1-k)(\overline{e}^{EAR} - \underline{e}^{EAR})\right],$$

which when combined with (6) gives (4).

Equation (4) reveals how the strength of the agent's incentives to insure himself against the compensation risk imposed on him by ex ante randomization varies with his preferences and with the parameters of the randomized contract. The smaller the cost difference between tasks (the smaller is λ) and the more risk-averse is the agent (the larger is r), the smaller the gap between the agent's optimal efforts on the two tasks, and as either $\lambda \to 1$ or $r \to \infty$, the optimal gap approaches zero. The smaller is the parameter k, the more different are the two possible compensation schedules, so the more risk the randomization per se imposes on the agent and hence the stronger 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. Similarly, as the incentive intensity on both tasks is scaled up by an increase in β, holding k fixed, the agent's incentive to self-insure rises, and the gap $\overline{e} - \underline{e}$ falls.

Recall that under a symmetric deterministic contract, the agent's effort choices change discontinuously as hidden information about the agent's preferences is introduced, i.e., as λ is increased from one: Efforts switch from being perfectly balanced at $\lambda = 1$ (the allocation preferred by the principal) to being completely focused efforts for any $\lambda > 1$. As a consequence, as we saw, the principal's payoff from a symmetric deterministic contract drops discontinuously as λ is raised from 1. Furthermore, even when the principal chooses the optimal deterministic scheme as a function of λ, as long as he tries to induce balanced efforts, his payoff drops discontinuously as λ is increased from 1. In contrast, under ex ante randomization, for any value of $k \in (-1, 1)$, both the agent's effort choices and the principal's payoff 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 the best deterministic scheme. EAR is also more robust to uncertainty about the magnitude of λ than is a deterministic scheme. If the principal tries to design an asymmetric deterministic menu to induce one type of agent to choose balanced efforts but is even slightly wrong about the magnitude of λ, her payoff will be discontinuously lower than if she were right. The performance of ex ante randomization does not display this extreme sensitivity.

Remark 1 *We have established Proposition 2 under the assumption that the principal can commit to randomizing half-half between the two compensation formulae.*[11] *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.*[12] *Thus the attractive properties of ex ante randomization are*

[11] 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.

[12] 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 half-half over the two mirror-image compensation schedules. Given that the principal randomizes half-half, the agents' optimal behavior

not crucially dependent on the principal's having the power to commit to the randomizing probability.

The effort-balancing incentives generated by EAR do, however, come at a cost in terms of the risk imposed on the risk-averse agent. In the principal's payoff expression (5), the two terms on the second line represent the total cost of the risk borne by the agent under EAR. 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, $w = \alpha + \beta x_2 + k\beta x_1$). Because, when $\lambda > 1$, the agent only partially insures himself against the risk imposed by the principal's randomization over compensation schedules, there is an additional component to the cost of risk borne by the agent, and this is represented by the final term on the second line in (5). Thus the total risk cost imposed by EAR exceeds that 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 payoff from EAR, it is helpful to define the variable $B \equiv \beta(1+k)$, because as equation (3) shows, aggregate effort $\overline{e} + \lambda \underline{e}$ is proportional to B. Using this definition and equations (3) and (4), we can re-express the principal's payoff (5) as a function of B and k:

$$\Pi^{EAR}(B,k) = \left(\underline{e} + \frac{\overline{e}}{\delta}\right) - \frac{B^2}{2\,(\lambda+1)^2} - \frac{1}{2} r\sigma^2 B^2 \left(\frac{1 + 2\rho k + k^2}{(1+k)^2}\right) - \frac{1}{2r} \ln\left(\frac{(\lambda+1)^2\,(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right). \quad (7)$$

Holding B fixed and varying k allows us to identify the effect of k on the principal's payoff from inducing any given level of aggregate effort. Equation (7) shows that increasing k has three effects. First, a larger k raises the gap between efforts on the tasks and, with B and hence aggregate effort $\overline{e}+\lambda\underline{e}$ held fixed, this larger gap lowers the principal's benefit $\underline{e}+\frac{\overline{e}}{\delta}$ whenever $\delta > \lambda$, i.e., whenever the principal's desire for balanced efforts is stronger than the agent's preference across tasks. Second, a larger k, because it induces the agent to choose less balanced efforts, raises the cost of compensating the agent for the risk imposed by the randomization per se.[13] This second effect of k also reduces the principal's payoff and is reflected in the final term in equation (7). 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 payoff, as reflected in the third term in equation (7).

In general, the optimal design of a contract with ex ante randomization 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. The following proposition describes how the optimal value of k varies with several parameters of the contracting environment.

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 the schedule rewarding task 1 more highly than task 2 with a probability greater than (less than) 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 choose the schedule rewarding task 2 more (less) highly than task 1.

[13] Holding the effort gap fixed, an increase in k would reduce the amount of risk imposed by the randomization per se, but when the agent optimally enlarges the effort gap in response to a rise in k, the overall cost of the risk imposed by the randomization rises.

Proposition 3 *For any given level of aggregate effort to be induced, the optimal level of k under EAR is smaller (the optimal weights on the performance measures should be more unequal)*
(i) the larger is δ (i.e., the stronger the principal's preference for balanced efforts);
(ii) the smaller is r, holding $r\sigma^2$ fixed (i.e., the less risk-averse the agent, holding fixed the importance of risk aversion under deterministic contracts);
(iii) the larger is ρ (i.e., the less scope for diversification of the risk from the shocks to measured performance).

It is also worth noting that the incentive instruments B (controlling aggregate effort) and k (controlling the degree of effort balance) are complements in the the principal's payoff function, in the sense that $\frac{\partial^2 \Pi}{\partial k \partial B} \geq 0$. Hence, ceteris paribus, when EAR provides strong incentives for aggregate effort, it is less valuable on the margin to lower the gap in efforts by lowering k.[14]

In Section 6, where we identify environments where ambiguous schemes outperform deterministic ones, we will build on these comparative statics results. For now, though, we turn to a second class of ambiguous contracts.

4.2 Ex Post Discretion

Under a contract involving *ex post discretion* (EPD), 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 $k \in (-1, 1)$. Just as under EAR, the agent is uncertain at the time he chooses his efforts whether his pay will be more sensitive to performance on task 1 or task 2, but with ex post discretion, unlike with EAR, the agent's choice of efforts can influence which compensation schedule is ultimately used. With EPD, as with EAR, the closer k is to 1, the more similar are the two possible compensation formulae, and if k were actually set equal to 1, EPD would involve no randomness at all and would collapse to the SD scheme.

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

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

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

Proposition 4 *(i) Under EPD, $k < \frac{1}{\lambda}$ is a necessary condition for each agent's optimal efforts on both tasks to be strictly positive.*
(ii) When for a given $k \in (-1, \frac{1}{\lambda})$, EPD induces interior solutions for the agents' effort choices,

[14] The larger is B, the smaller is the gap in efforts for any given k, and therefore the lower the marginal benefit to reducing k in order to further lower the effort gap; this interaction is reflected in the second term in (7). In addition, the larger is B, the larger is the risk imposed by the shocks to measured performance, so the larger is the marginal benefit to improving diversification by increasing k; this interaction is reflected in the fourth term in (7).

each type of agent chooses effort on his less-costly task, $\overline{e}^{EPD}$, and effort on his more-costly task, $\underline{e}^{EPD}$, satisfying

$$\overline{e}^{EPD} + \lambda \underline{e}^{EPD} = \frac{\beta(1+k)}{\lambda+1} \tag{8}$$

$$\frac{\lambda-k}{1-k\lambda} = \exp\left[r\beta(1-k)(\overline{e}^{EPD}-\underline{e}^{EPD})\right] \frac{\Phi\left(\frac{(1-k)(\overline{e}^{EPD}-\underline{e}^{EPD})+r(\sigma^k)^2\beta\left(1-\rho^k\right)}{\theta^k}\right)}{\Phi\left(\frac{-(1-k)(\overline{e}^{EPD}-\underline{e}^{EPD})+r(\sigma^k)^2\beta(1-\rho^k)}{\theta^k}\right)}, \tag{9}$$

where $(\sigma^k)^2 \equiv var(x_1 + kx_2) = \sigma^2(1 + 2\rho k + k^2)$*,* $\rho^k \equiv corr(x_1 + kx_2, x_2 + kx_1) = \frac{\rho+2k+\rho k^2}{1+2\rho k+k^2}$*, and* $\theta^k \equiv \sigma^k[2(1-\rho^k)]^{\frac{1}{2}}$*.*

(iii) When for a given $\lambda > 1$ and $k \in (-1, \frac{1}{\lambda})$, EPD and EAR both induce interior solutions for efforts, then $\overline{e}^{EPD} - \underline{e}^{EPD} < \overline{e}^{EAR} - \underline{e}^{EAR}$.

(iv) The gap in efforts, $\overline{e}^{EPD} - \underline{e}^{EPD}$, is increasing in λ, approaching 0 as $\lambda \to 1$; decreasing in r, approaching 0 as $r \to \infty$; increasing in $\sigma^2(1-\rho)$, approaching 0 as $\sigma^2(1-\rho) \to 0$; and increasing in k, approaching 0 as $k \to -1^+$.

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

$$\begin{aligned}\Pi^{EPD}(\beta,k) = {} & \underline{e}^{EPD} + \frac{1}{\delta}\overline{e}^{EPD} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^k)^2\beta^2 \\ & -\frac{1}{r}\ln\left[\exp\{-r\beta(1-k)(\overline{e}^{EPD}-\underline{e}^{EPD})\}\Phi(-)+\Phi(+)\right] \\ & -\beta(1-k)(\overline{e}^{EPD}-\underline{e}^{EPD})\Phi\left(-\frac{(1-k)(\overline{e}^{EPD}-\underline{e}^{EPD})}{\theta^k}\right) \\ & +\beta\theta^k\phi\left(\frac{(1-k)(\overline{e}^{EPD}-\underline{e}^{EPD})}{\theta^k}\right).\end{aligned} \tag{10}$$

where

$$\Phi(-) \equiv \Phi\left(\frac{-(1-k)(\overline{e}^{EPD}-\underline{e}^{EPD})+r(\sigma^k)^2\beta(1-\rho^k)}{\theta^k}\right)$$

$$\Phi(+) \equiv \Phi\left(\frac{(1-k)(\overline{e}^{EPD}-\underline{e}^{EPD})+r(\sigma^k)^2\beta\left(1-\rho^k\right)}{\theta^k}\right)$$

The first part of the proposition holds for the same reason as the first part of Proposition 2: if $k \geq \frac{1}{\lambda}$, then *whichever of the two compensation schemes is ultimately chosen by the principal*, the ratio of marginal return to marginal cost would be 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 it would be optimal for both types of agent to exert effort only on their preferred task.

Proposition 4 shows that aggregate effort, $\overline{e} + \lambda\underline{e}$, is the same under EPD as under EAR—compare equations (8) and (3). Since both schemes are certain to reward one task at rate β and the other at rate $k\beta$, the sum of the expected marginal returns to effort on the two tasks is $(1+k)\beta$ under both schemes, 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 then imply

$$\frac{\lambda - k}{1 - k\lambda} = \frac{E\left[U'(\cdot) I_{\{\underline{x} \text{ is rewarded more highly than } \overline{x}\}}\right]}{E\left[U'(\cdot) I_{\{\overline{x} \text{ is rewarded more highly than } \underline{x}\}}\right]}, \tag{11}$$

but for ex post discretion

$$\begin{aligned} &\frac{E\left[U'(\cdot) I_{\{\underline{x} \text{ is rewarded more highly than } \overline{x}\}}\right]}{E\left[U'(\cdot) I_{\{\overline{x} \text{ is rewarded more highly than } \underline{x}\}}\right]} \\ &\quad = \exp\left[r\beta(1-k)(\overline{e}^{EPD} - \underline{e}^{EPD})\right] \frac{\Phi\left(\frac{(1-k)(\overline{e}^{EPD} - \underline{e}^{EPD}) + r(\sigma^k)^2 \beta (1-\rho^k)}{\theta^k}\right)}{\Phi\left(\frac{-(1-k)(\overline{e}^{EPD} - \underline{e}^{EPD}) + r(\sigma^k)^2 \beta (1-\rho^k)}{\theta^k}\right)}, \end{aligned}$$

which when combined with equation (11) gives us equation (9).

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 choice of compensation schedule. Under EPD, the insurance motive is still present, because at the time the agent chooses efforts, he is uncertain about which compensation schedule the principal will select. Now, though, there is an additional incentive for the agent to balance his efforts: 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 less likely that task is to be the more highly compensated one, so the lower the relative marginal return to that task. Formally, the right-hand side of equation (9), which is increasing in $\overline{e} - \underline{e}$, is strictly greater than the right-hand side of equation (4) for all $\overline{e} - \underline{e} > 0$, and this implies

$$\overline{e}^{EPD} - \underline{e}^{EPD} < \overline{e}^{EAR} - \underline{e}^{EAR} \qquad \forall \lambda > 1. \tag{12}$$

Intuitively, we might expect that the principal's freedom, under EPD, to choose the compensation schedule 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 EPD than under EAR, as follows from (8) and (12). Nevertheless, aggregate effort $\overline{e} + \lambda\underline{e}$, and hence the costs of effort incurred, are the same under the two schemes.

Equation (9) allows us to show that under EPD, the agent optimally chooses a gap between the effort levels on the two tasks that is smaller the smaller is λ (because it is less costly for the agent to choose balanced efforts) and the larger is r (because the stronger desire to self-insure is the dominant effect), and as either $\lambda \to 1$ or $r \to \infty$, this gap approaches zero. These results parallel those for EAR. However, while σ^2 and ρ have no effect on the gap in effort levels under EAR, under EPD 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 compensation schedule the principal chooses, so gives the agent a stronger incentive to balance his efforts. As $\sigma^2(1-\rho) \to 0$, for example because the shocks become perfectly correlated, optimal efforts become perfectly balanced.

Under EPD, just as under EAR, reducing the parameter k, and so making the two possible compensation schedules more different, induces the agent to choose more balanced efforts. While the effect of k on the effort gap $\overline{e} - \underline{e}$ is more complex under EPD than under EAR, nevertheless the proof of part 4. of Proposition 4 shows that the effects of k that operate under EAR are the dominant ones under EPD. As a consequence, under EPD, just as under EAR, the optimal level of k, for any given level of aggregate effort to be induced, will be smaller, the stronger is the principal's preference for balanced efforts (i.e., the larger is the complementarity parameter δ).

What is the cost of the risk imposed by ex post discretion on the risk-averse agent, and how does it compare to that imposed by ex ante randomization? In the principal's payoff expression (10), the total cost of the risk imposed is given by $-\frac{1}{2}r(\sigma^g)^2\beta^2$ plus the terms on the final three lines. We can show that, in fact, EPD with coefficients β and k imposes *lower* risk costs than would either of the deterministic contracts $w = \alpha + \beta x_1 + k\beta x_2$ or $w = \alpha + \beta x_2 + k\beta x_1$, which would impose risk cost $-\frac{1}{2}r(\sigma^g)^2\beta^2$. Formally, this claim corresponds to the result that the principal's payoff in (10) is *greater* than $\underline{e}^{EPD} + \frac{1}{\delta}\overline{e}^{EPD} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^g)^2\beta^2$. The intuitive reason why this result holds is that the variance of the wage under EPD, $w = \min\{\alpha + \beta x_1 + k\beta x_2, \alpha + \beta x_2 + k\beta x_1\}$, is *lower* than the variance of either $\alpha + \beta x_1 + k\beta x_2$ or $\alpha + \beta x_2 + k\beta x_1$.[15] Section 4.1 showed, by contrast, that for any given β and k, EAR imposes *higher* risk costs than would either of the deterministic contracts above.

4.3 Ex Ante Randomization versus Ex Post Discretion

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

Proposition 5 *If, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$, both EAR and EPD induce interior solutions for the agents' effort choices, and if $\delta \geq \lambda$, then EPD generates at least as great a payoff for the principal as EAR.*

The condition $\delta \geq \lambda$ ensures that the smaller gap in efforts under ex post discretion, 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.[16]

[15] We can also show that the risk cost imposed by EPD is increasing in the gap $\overline{e}^{EPD} - \underline{e}^{EPD}$, reflecting the fact that the variance of $w = \min\{\alpha + \beta x_1 + k\beta x_2, \alpha + \beta x_2 + k\beta x_1\}$ is increasing in $\overline{e}^{EPD} - \underline{e}^{EPD}$.

[16] It is natural to wonder whether other types of ambiguous incentive schemes can be more attractive than ex ante randomization and ex post discretion. Suppose, for example, that the principal implements the following more general form of ambiguous scheme: With probability $q \in [0, 1]$, he waits until he has observed x_1 and x_2 and then chooses whether to pay the agent $\alpha + \beta x_1 + k\beta x_2$ or $\alpha + \beta x_2 + k\beta x_1$, while with probability $1 - q$ he randomizes ex ante with equal probabilities between the two schedules $w = \alpha + \beta x_1 + k\beta x_2$ and $w = \alpha + \beta x_2 + k\beta x_1$. Clearly, EAR corresponds to the special case where $q = 0$ and EPD to that where $q = 1$. We can show that for any given (q, β, k) that induce interior effort choices from the agent, an increase in q leaves aggregate effort $\overline{e} + \lambda\underline{e}$ unchanged, reduces the gap in efforts $\overline{e} - \underline{e}$, and reduces the cost of risk imposed on the agent. Therefore, as long as $\delta > \lambda$, the principal's payoff is increased when q increases, so the optimal ambiguous scheme within this family is what we have termed EPD.

In supporting online material, we analyze randomized incentive schemes involving menus and show that such menus bring no additional benefit. Hence, for the comparisons in the following two sections, the only ambiguous incentive schemes we consider are EAR and EPD.

5 When Are Deterministic Contracts Optimal?

This section identifies three environments in which both types ambiguous 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 an ambiguous 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 ambiguous 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 6 *For any given $\beta > 0$ and $k \in (-1, 1)$, both EAR and EPD yield lower payoff for the principal than a suitably designed symmetric deterministic (SD) scheme, if any of the following conditions holds:*
(i) $\lambda = 1$;
(ii) EAR and EPD induce the agent to exert effort only on his preferred task;
(iii) $\delta \leq \lambda$.

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 EAR or EPD. 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 EAR and EPD have the potential to induce better-balanced efforts. In general, therefore, the principal faces a trade-off in choosing between ambiguous and deterministic incentive schemes. Ambiguous 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 6 are ones under which this trade-off does not in fact arise. Under condition 1 or 2, ambiguous schemes are no better than a SD scheme at inducing balanced efforts: in the former case, the SD scheme, like the ambiguous schemes, induces perfectly balanced efforts, and in the latter case, even the ambiguous schemes induce corner solutions. Under condition 3, the socially efficient effort allocation involves fully focused efforts. Hence, for any fixed level of aggregate effort and hence fixed cost of efforts incurred, a shift towards more balanced efforts would actually reduce (at least weakly) the principal's payoff. Therefore, under any of conditions 1, 2, or 3, the potential benefits of ambiguous schemes in inducing better-balanced efforts do not actually materialize, and the principal's optimal incentive scheme is a deterministic one.

Proposition 6 has an informative corollary:

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

Recall that the principal's payoff from 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 ambiguous contracts depends also on the individual values of r and σ^2. Proposition 2 shows that as r falls, the gap in efforts under ex ante randomization rises (because the agent's desire for self-insurance diminishes), and Proposition 3 shows that as r falls and σ^2 rises, both changes lead to a larger gap in efforts under ex post discretion, 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 $r\sigma^2$, ambiguous schemes perform badly relative to deterministic ones when r is very small and σ^2 is very large, because in such settings, ambiguous schemes generate extremely weak incentives to choose balanced efforts.

6 When Are Ambiguous Contracts Optimal?

We now identify three environments in which ambiguous contracts, when designed optimally, can be shown to dominate the best deterministic scheme. In each of these environments, EAR and EPD, with the parameter k adjusted optimally, both induce the agent to choose *perfectly* balanced efforts, and EAR is as attractive for the principal as EPD. 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 optimally designed EAR and EPD generate a payoff for the principal arbitrarily close to that she could achieve if she knew the agent's preferences across tasks, so ambiguous schemes eliminate the efficiency losses from the agent's hidden information. In these three environments, the optimal incentive scheme is an ambiguous one whenever, in the "no hidden information benchmark", it would be optimal to induce balanced efforts, i.e., whenever the principal's complementarity parameter δ exceeds δ^{NHI} (as defined in (1)). If $\delta < \delta^{NHI}$, then it is optimal for the principal to use a symmetric deterministic menu (SDM) and abandon any attempt to induce balanced efforts. An asymmetric deterministic menu (ADM) is never optimal in these environments.

6.1 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 $\lambda \to 1^+$.

For ex ante randomization recall that the agent's effort choices and the principal's payoff are continuous in λ at $\lambda = 1$. Proposition 2 established that as $\lambda \to 1$, $\overline{e}^{EAR} - \underline{e}^{EAR} \to 0$ for any $k \in (-1, 1)$. Equation (7) shows how varying k affects the principal's payoff from EAR, holding fixed the level of aggregate effort induced. Whereas in general, as discussed in Section 4.1, increasing k has opposing effects on the principal's payoff, in the limit as $\lambda \to 1$, the situation is dramatically simpler. As $\lambda \to 1$, equation (7) becomes

$$\Pi^{EAR}(B,k) = \frac{(\delta+1)}{\delta}\frac{B}{4} - \frac{B^2}{8} - \frac{1}{2}r\sigma^2 B^2\left(\frac{1+2\rho k+k^2}{(1+k)^2}\right). \tag{13}$$

With perfectly balanced efforts ensured by λ approaching 1, an increase in k has only one effect on the principal's payoff from inducing any given level of aggregate effort: it improves the diversification of the risk from the shocks to measured performance, as reflected in the final term of equation (13). Hence, as $\lambda \to 1$, the principal's payoff from EAR 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 payoff from EAR 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 payoff approaches

$$\Pi^{EAR}(B) = \frac{(\delta+1)}{\delta}\frac{B}{4} - \frac{B^2}{8} - \frac{1}{4}r\sigma^2 B^2(1+\rho). \tag{14}$$

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

$$\frac{(\delta+1)^2}{8\delta^2(1+2r\sigma^2(1+\rho))}, \tag{15}$$

which is the payoff the principal would achieve, *at* $\lambda = 1$, from a SD contract. (Recall equation (29).) This is also the payoff the principal would achieve in the "no hidden information benchmark" from the contract pair (C_1^{bal}, C_2^{bal}) in the limit as $\lambda \to 1$.[17]

For ex post discretion, too, as $\lambda \to 1$, $\overline{e}^{EPD} - \underline{e}^{EPD} \to 0$ for any $k \in (-1, 1)$, as follows from the fact that the gap in efforts on the two tasks is smaller under EPD than under EAR. We now convert the payoff expression (10) for EPD given in Proposition 4 into one as a function of $B \equiv \beta(1+k)$ and k, as we did for EAR, and simplify it in the limit as $\lambda \to 1$. This yields

$$\begin{aligned}\Pi^{EPD}(B,k) = \frac{(\delta+1)}{\delta}\frac{B}{4} \quad &- \quad \frac{B^2}{8} - \frac{1}{2}r\sigma^2 B^2\left(\frac{1+2\rho k+k^2}{(1+k)^2}\right) \\ &+ \quad \frac{1}{r}\left\{\frac{rB\theta^g}{1+k}\phi(0) - \ln\left[2\Phi\left(\frac{rB\theta^g}{2(1+k)}\right)\right]\right\}\end{aligned} \tag{16}$$

It can be shown that this payoff expression, like that for EAR, 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

[17] Recall that the payoff in the NHI benchmark is continuous as $\lambda \to 1$, in contrast to the discontinuity at $\lambda = 1$ under hidden information.

correlation between $x_1 + kx_2$ and $x_2 + kx_1$, increases. Therefore, as $\lambda \to 1$, the principal's payoff from EPD 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 (16) approach 0 since $\rho^k \equiv corr(x_1 + kx_2, x_2 + kx_1) \to 1$ and hence $\theta^k \equiv \sigma^k[2(1-\rho^k)]^{\frac{1}{2}} \to 0$. Hence, in the limit as $\lambda \to 1$, the principal's maximized payoff from EPD with k adjusted optimally also approaches the level he would achieve, at $\lambda = 1$, from a SD contract.

When the principal is restricted to deterministic incentive schemes, then as shown in Section 3, even as λ gets arbitrarily close to 1 ($\lambda \to 1^+$), it is impossible for the principal to induce balanced efforts from more than one type of agent. Consequently, whenever $\delta > \lim_{\lambda \to 1} \delta^{NHI}(\lambda, r\sigma^2, \rho)$, the payoff from an asymmetric deterministic menu, even in the limit as $\lambda \to 1^+$, remains strictly below that which would be achievable from a SD contract at $\lambda = 1$.

Proposition 7 *Consider the limiting case as $\lambda \to 1^+$. Under both EAR and EPD, for any given level of aggregate effort, $\overline{e} + \lambda \underline{e}$, to be induced:*
(i) the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any $k \in (-1, 1)$;
(ii) the optimal value of $k \to 1^-$;
*(iii) with k adjusted optimally, the principal's payoff under both ambiguous schemes approaches her payoff in the NHI benchmark from (C_1^{bal}, C_2^{bal}) as $\lambda \to 1$, which equals her payoff from the symmetric deterministic (SD) contract **at** $\lambda = 1$.*
Therefore, for $\delta > \lim_{\lambda \to 1} \delta^{NHI}(\lambda, r\sigma^2, \rho)$, EAR and EPD with k adjusted optimally dominate the best deterministic scheme under hidden information. For $\delta < \lim_{\lambda \to 1} \delta^{NHI}(\lambda, r\sigma^2, \rho)$, the principal's optimal incentive scheme is a symmetric deterministic menu (SDM).

6.2 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 EAR and EPD 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 (4) in Proposition 2 shows that, for any $k \in (1, \frac{1}{\lambda})$, EAR induces the agent to choose perfectly balanced efforts: $\overline{e}^{EAR} - \underline{e}^{EAR} = 0$. 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 4 showed that $\overline{e}^{EPD} - \underline{e}^{EPD} \leq \overline{e}^{EAR} - \underline{e}^{EAR}$, it follows that in this limiting case, EPD 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 payoff, as well as the efforts induced, under all deterministic schemes remain the same, when risk aversion becomes very large and exogenous shocks very small, both types of ambiguous incentive scheme generate very strong incentives to choose balanced efforts.

In the limit as $r \to \infty$ and $\sigma^2 \to 0$, equation (7) becomes

$$\Pi^{EAR}(B,k) = \frac{(\delta+1)}{\delta}\frac{B}{(\lambda+1)^2} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2}RB^2\left(\frac{1+2\rho k+k^2}{(1+k)^2}\right). \tag{17}$$

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 payoff from inducing any given level of aggregate effort: it improves the diversification of the risk imposed on the agent from the shocks to measured performance. Hence, in this limiting case, too, the principal's payoff from EAR 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 payoff expression in (17) approaches

$$\Pi^{EAR}(B) = \frac{(\delta+1)}{\delta}\frac{B}{(\lambda+1)^2} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2}RB^2\left(\frac{\lambda^2+2\rho\lambda+1}{(\lambda+1)^2}\right). \tag{18}$$

This is exactly the payoff the principal would obtain, in the "no hidden information benchmark", from using (C_1^{bal}, C_2^{bal}) to induce perfectly balanced efforts and setting $\beta = \frac{B}{1+\lambda}$ (see equation (32)).

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

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

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

Proposition 8 *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 EAR and EPD, for any given level of aggregate effort, $\overline{e} + \lambda\underline{e}$, to be induced:*
(i) the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any λ and for any $k < \frac{1}{\lambda}$;
(ii) the optimal value of $k \to \left(\frac{1}{\lambda}\right)^-$;
(iii) with k adjusted optimally, the principal's payoff under both ambiguous schemes approaches her payoff in the NHI benchmark from (C_1^{bal}, C_2^{bal}) with incentive coefficient β chosen to induce the same level of aggregate effort.
Therefore, for $\delta > \delta^{NHI}(\lambda, R, \rho)$, EAR and EPD with k adjusted optimally dominate the best deterministic scheme under hidden information. For $\delta < \delta^{NHI}(\lambda, R, \rho)$, the principal's optimal incentive scheme is a symmetric deterministic menu (SDM).

6.3 The Limiting Case of Perfect Correlation of the Shocks

Under ex post discretion, 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 EPD, as long as $k < \frac{1}{\lambda}$, the agent's optimal efforts on the two tasks will be equal. To see this, consider an 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., $e_1 + \lambda e_2 = e^* + \lambda e^*$. Both the agent's cost of effort and the risk premium are the same under the two effort pairs. Therefore, the only effect of switching from (e_1, e_2) to (e^*, e^*) on the agent's expected utility is the effect on the expected wage. For both effort pairs, since $\rho = 1$, the agent will receive $\alpha + \beta(x_2 + kx_1)$, so the switch causes the expected wage to change by

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

and this expression is strictly positive whenever $k < \frac{1}{\lambda}$. If, instead, a type-1 agent switched from (e_1, e_2) with $e_1 < e_2$ to (e^*, e^*) such that aggregate effort was unchanged, the switch 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 EPD.

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 $\alpha + \beta(1+k)e + \beta(1+k)\epsilon$, where $\epsilon \sim N(0, \sigma^2)$, that is, ϵ 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 payoff under EPD for a given β and $k \in (-1, \frac{1}{\lambda})$ is

$$\Pi^{EPD}(\beta, k) = \frac{\delta+1}{\delta}\frac{\beta(1+k)}{(\lambda+1)^2} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2(1+k)^2\beta^2. \tag{19}$$

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

$$\Pi^{EPD}(B, k) = \frac{\delta+1}{\delta}\frac{B}{(\lambda+1)^2} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2 B^2. \tag{20}$$

This payoff expression 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. Therefore, under EPD, when $\rho = 1$, any value of $k \in (-1, \frac{1}{\lambda})$ is optimal. Note that the payoff expression (20) matches what the principal would obtain, in the "no hidden information benchmark" with $\rho = 1$, from using (C_1^{bal}, C_2^{bal}) to induce perfectly balanced efforts and setting $\beta = \frac{B}{1+\lambda}$ (see equation (32)). Thus, in this limiting environment as well, EPD allows the principal to achieve a payoff as high as she would achieve in the absence of hidden information.

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

$$\Pi^{EAR}(B,k) = \frac{(\delta+1)B}{\delta(\lambda+1)^2} - \frac{(\delta-\lambda)\ln\left(\frac{\lambda-k}{1-k\lambda}\right)}{\delta(\lambda+1)rB\left(\frac{1-k}{1+k}\right)} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2B^2 - \frac{1}{2r}\ln\left(\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right). \quad (21)$$

As with EPD, since $\rho = 1$, varying k has no effect on the diversification of the risk from the shocks to performance. Consequently, the only effects of increasing k, holding aggregate effort fixed, are the negative ones 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 (21)) and also results in the agent bearing more risk from the randomization (the final term in (21)). With $\rho = 1$, it is therefore optimal under EAR, 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 2), and the principal achieves a payoff arbitrarily close to[18]

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

This is the same payoff expression as arose under EPD for any value of $k \in (-1, \frac{1}{\lambda})$.

Proposition 9 *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:*
(i) under EPD, 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;
(ii) under EAR, the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any λ as $k \to -1^+$, and the optimal value of $k \to -1^+$;
(iii) with k adjusted optimally, the principal's payoff under both ambiguous schemes approaches her payoff in the NHI benchmark from (C_1^{bal}, C_2^{bal}) with incentive coefficient β chosen to induce the same level of aggregate effort.
Therefore, for $\delta > \delta^{NHI}(\lambda, r\sigma^2, 1)$, EAR and EPD with k adjusted optimally dominate the best deterministic scheme under hidden information. For $\delta < \delta^{NHI}(\lambda, r\sigma^2, 1)$, the principal's optimal incentive scheme is a symmetric deterministic menu (SDM).

6.4 Discussion

We have identified three environments in which ambiguous contracts, when designed optimally, dominate the best deterministic scheme. In all three environments, EAR and EPD, with the para-

[18] 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.

meter k adjusted optimally, both induce both types of agent to choose perfectly balanced efforts, and EAR is as profitable for the principal as EPD. Both types of ambiguous scheme generate a payoff for the principal arbitrarily close to her payoff in the "no hidden information benchmark", so in these environments, ambiguous incentive schemes eliminate the efficiency losses from the agent's better knowledge of the environment. Figure 2 summarizes our findings on the optimal contract choice for the three limiting settings.

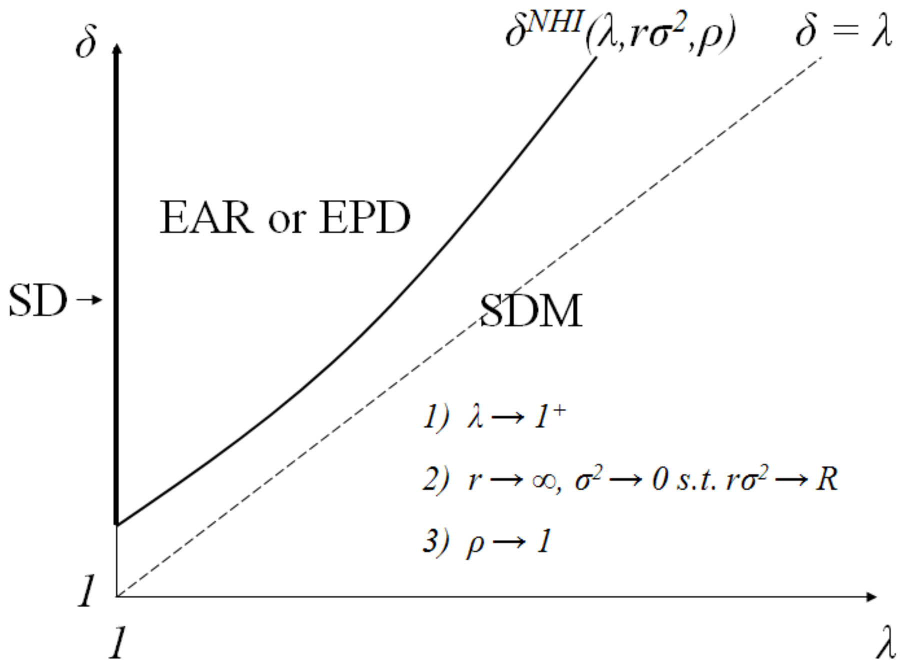

Figure 2: Optimal incentive schemes in 3 limiting settings

In each of these settings, optimally designed ambiguous schemes are preferable to the best deterministic scheme whenever, in the NHI benchmark, the principal wants to induce balanced efforts, that is, whenever her complementarity parameter δ exceeds $\delta^{NHI}(\lambda, r\sigma^2, \rho)$. If, in the NHI benchmark, the risk premium incurred in inducing balanced efforts would be too large relative to the benefits of balance, then in these settings, ambiguous schemes are dominated by a SDM, which induces both types of agent to focus their efforts on their preferred task.

In none of these settings is the asymmetric deterministic menu (ADM) ever strictly the most attractive 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 ambiguous 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. Whenever $\lambda > 1$, these rents make the ADM strictly less profitable, for all values of δ, than the best alternative scheme.

The comparative statics results for δ^{NHI} discussed after equation (7) tell us when, in the three environments studied above, ambiguous schemes are more likely to dominate the best deterministic scheme under hidden information. Ambiguous schemes are more likely to be optimal the smaller is $r\sigma^2$: this reflects the imposition of greater risk on the agent by EAR/EPD than SDM per unit of aggregate effort induced. In the limit as $r\sigma^2 \to 0$, ambiguous schemes are optimal in these environments if and only if $\delta > \lambda$: with risk no longer having a cost in this limit, the principal prefers ambiguous to deterministic schemes according to whether or not the perfectly balanced efforts induced by the ambiguous schemes are socially efficient. In the environments of Sections 6.2 and 6.3, ambiguous schemes are more likely to be optimal the smaller is 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 the ambiguous schemes and the SDM widens. Finally, in the environments of Sections 6.1 and 6.2, ambiguous schemes are more likely to be preferred the smaller is ρ, since a smaller ρ both i) reduces the gains from optimal insurance offered by the SDM and ii) reduces the extra risk costs under the ambiguous schemes of basing compensation on both performance measures.

7 Robustness and Extensions

7.1 Ex Ante Randomization and the Choice of How Many Tasks to Reward

We have assumed so far that the job performed by the agent has only two distinct dimensions (tasks) and that noisy measures of performance on both tasks are used in randomized incentive schemes. When, however, performance on a job has many distinct dimensions, the costs of monitoring the different dimensions may become significant. In such settings, the principal can economize on monitoring costs, while still providing incentives for balanced efforts, by randomizing over compensation schedules each of which rewards only a subset of dimensions of performance. We now study some of the tradeoffs involved in the design of randomized incentive schemes in environments with many tasks.

Let the job performed by the agent consist of $n > 2$ tasks, for each of which measured performance $x_j = e_j + \epsilon_j$, where $(\epsilon_1, \ldots, \epsilon_n)$ have a symmetric multivariate normal distribution with mean 0, variance σ^2, and pairwise correlation $\rho \geq 0$. Suppose there are n equally likely types of agent, with the agent of type i having cost function $c_i(e_1, \ldots, e_n) = \frac{1}{2}(\lambda e_i + \sum_{j \neq i} e_j)^2$. Thus each type of agent has a particular dislike for exactly one of the n tasks, and λ measures the intensity of this dislike. Let the principal's benefit function be

$$B(e_1, \ldots, e_n) = \min\{e_1, \ldots, e_n\} + \frac{1}{\delta}\left(\sum_{j=1}^{n} e_j - \min\{e_1, \ldots, e_n\}\right),$$

where as before δ parameterizes the strength of the principal's desire for a balanced profile of efforts across tasks. With the cost and benefit functions specified, the socially efficient profile of efforts is perfectly balanced ($e_i = e_j$ for all $i \neq j$) whenever $\delta > \lambda$.

We will study the following family of incentive schemes with ex ante randomization, parameterized by κ, the number of tasks rewarded: Each subset of κ out of n tasks is chosen with equal probability, and each task in the chosen subset is rewarded at rate β; whichever subset is chosen, the lump-sum payment is α. We will not explicitly model the direct costs of generating the performance measures. Instead we will focus on the incentive and risk costs of varying the number of tasks κ included in each of the possible compensation schedules.

Denote by $\underline{e}$ each type of agent's effort on his disliked task and by $\overline{e}$ his effort on each of the other tasks. If, for a given κ, the agent's optimal effort choices are interior, then aggregate effort $(\lambda\underline{e} + (n-1)\overline{e})$ and the gap in efforts $\overline{e} - \underline{e}$ satisfy, respectively,

$$\lambda\underline{e} + (n-1)\overline{e} = \frac{\kappa\beta}{n-1+\lambda} \qquad \text{and} \qquad \overline{e} - \underline{e} = \frac{1}{r\beta}\ln\left[\frac{\lambda(n-\kappa)}{(n-1)-(\kappa-1)\lambda}\right]. \tag{22}$$

Reducing κ, the number of tasks rewarded, reduces the pairwise correlation between the different levels of compensation the agent might receive under the randomized scheme and so gives the agent a stronger incentive to self-insure. As a result, the the agent's optimal effort profile is more balanced ($\overline{e} - \underline{e}$ is smaller), the smaller is the number of tasks rewarded.

Since aggregate effort is proportional to $\kappa\beta$, define $\tilde{\beta} \equiv \kappa\beta$. Using equations (22), we can write the principal's payoff as a function of $\tilde{\beta}$, κ, and n, given that the fixed payment α is set to ensure zero rents:

$$\begin{aligned} \Pi(\tilde{\beta}, \kappa, n) \quad &= \quad \left(\underline{e} + \frac{(n-1)}{\delta}\overline{e}\right) - \frac{\tilde{\beta}^2}{2\,(n-1+\lambda)^2} \\ &\qquad -\frac{1}{2}r\sigma^2\tilde{\beta}^2\left(\frac{1+\rho(\kappa-1)}{\kappa}\right) - \frac{1}{nr}\ln\left[\frac{(n-\kappa)^{n-\kappa}(n-1+\lambda)^n}{n^n(\lambda)^\kappa\left((n-1)-(\kappa-1)\lambda\right)^{n-\kappa}}\right], \end{aligned} \tag{23}$$

where

$$\underline{e} + \frac{(n-1)}{\delta}\overline{e} = \left(\frac{\delta+n-1}{\delta}\right)\frac{\tilde{\beta}}{(n-1+\lambda)^2} - \frac{(\delta-\lambda)(n-1)\kappa}{\delta(n-1+\lambda)r\tilde{\beta}}\ln\left[\frac{\lambda(n-\kappa)}{(n-1)-(\kappa-1)\lambda}\right]. \tag{24}$$

Holding $\tilde{\beta}$ and n fixed and varying κ allows us to isolate the effect of changing the number of tasks rewarded, holding fixed the level of aggregate effort induced. Comparison of equations (23)-(24) with equation (7) reveals that changes in κ have qualitatively the same three effects on the principal's payoff in this n-task model as do variations in the weighting coefficient k in EAR in the original two-task model. Specifically, an increase in κ, because it raises the gap between $\overline{e}$ and $\underline{e}$, has two negative effects: i) it lowers the principal's benefit $\underline{e} + \frac{(n-1)}{\delta}\overline{e}$ when aggregate effort is held fixed, as long as $\delta > \lambda$ (see (24)), and ii) it raises the cost of compensating the agent for the risk imposed by the random choice of which subset of tasks to reward (see the final term in (23)). At the same time, raising κ also has the positive effect of improving the diversification of the risk stemming from the shocks to measured performance, thereby reducing the risk premium per unit of aggregate effort induced (see the second-to-last term in (23)).

Given the observations above on the qualitative similarity of varying κ in the n-task model and k in the two-task model, we can conclude:

Proposition 10 *Consider the model with n tasks and a completely random choice of which subset of κ of them to reward. For any given level of aggregate effort to be induced, the optimal number of tasks to reward is smaller*
(i) the larger is δ (i.e., the stronger the principal's preference for balanced efforts);
(ii) the smaller is r, holding $r\sigma^2$ fixed (i.e., the less risk-averse the agent, holding fixed the importance of risk aversion under deterministic contracts);
(iii) the larger is ρ (i.e., the less scope for diversification of the risk from the shocks to measured performance).

7.2 Imperfect Substitutability of Efforts for the Agent

The analysis so far has focused on the case where efforts are perfect substitutes in the agent's cost function. This assumption was made for convenience: although it does not qualitatively affect the performance of ex ante randomization and ex post discretion, it simplifies the characterization of the optimal linear deterministic incentive scheme. We now show that our key findings continue to hold even when we introduce imperfect substitutability of efforts for the agent. Specifically, it remains true that i) if tasks are sufficiently complementary for the principal, ambiguous schemes are superior to deterministic schemes in settings where they generate very strong incentives for balanced efforts and ii) in such settings, ambiguous schemes completely eliminate the efficiency losses from the agent's better knowledge of the environment.

Let the two equally likely types of agent have cost functions of the form

$$c(\overline{e}, \underline{e}) = \frac{1}{2}\left(\overline{e}^2 + 2s\lambda\overline{e}\underline{e} + \lambda^2\underline{e}^2\right) \tag{25}$$

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. The parameter λ continues to measure each type's degree of bias towards his preferred task, and the new parameter $s \in [0, 1]$ measures the degree of substitutability of efforts. Note that $s = 1$ represents perfect substitutability and $s = 0$ no substitutability. We will focus on the case where $s \geq \frac{1}{\lambda}$, which represents a situation of high, but imperfect, substitutability. For simplicity, we will also specialize to the case where the principal's preference for balanced efforts is strongest, that is, $B(e_1, e_2) = \min\{e_1, e_2\}$, which corresponds to $\delta \to \infty$.

In the "no hidden information benchmark" setting, the principal will offer each type of agent a contract of the form $w = \alpha + \beta\overline{x} + v\beta\underline{x}$ with $v \geq 1$, where $\overline{x}$ denotes measured performance on the agent's preferred task and $\underline{x}$ measured performance on the other task. The parameter v determining the relative rate of reward on the two tasks is a choice variable for the principal, and under the assumptions above, the optimal choice of v, v^{NHI}, can be shown to induce each type of agent to choose equal efforts on the two tasks: $v^{NHI} = \frac{\lambda(\lambda+s)}{1+s\lambda}$. The principal's payoff in the NHI benchmark is continuous in λ and s. In the limit as $\lambda \to 1$ and $s \to 1$, $v^{NHI} \to 1$: the optimal contract for each type approaches the SD contract $w = \alpha + \beta x_1 + \beta x_2$, and the principal's payoff in the NHI benchmark approaches her payoff from the SD contract at $\lambda = 1$ (as given by equation (2)).

When the agent is privately informed about his preferences across tasks and $s \geq \frac{1}{\lambda}$, then just as in the original model, it is not possible with a menu of deterministic linear contracts to induce both types of agent to choose strictly positive effort on both tasks. The optimal deterministic menu will be an asymmetric deterministic menu which induces one type to choose (relatively) balanced efforts while the other type chooses fully focused efforts. It is easy to derive an upper bound on the principal's payoff from an ADM and to show that this bound is strictly less than her payoff in the NHI benchmark. Furthermore, whereas the payoff in the NHI benchmark approaches the value in equation (2) as $\lambda \to 1^+$ and $s \to 1^-$, the payoff from an ADM is bounded away from this level.

Ex ante randomization continues to give the risk-averse agent an incentive to partially self-insure by choosing relatively balanced efforts on the two tasks, and ex post discretion continues to give even stronger incentives for balance because of the agent's ability to influence, through his efforts, which task is more highly rewarded. Interior optimal efforts under EAR satisfy

$$\beta(1+k) = \frac{\partial c}{\partial \overline{e}} + \frac{\partial c}{\partial \underline{e}} \tag{26}$$

and

$$\frac{\frac{c_2}{c_1} - k}{1 - k\frac{c_2}{c_1}} = \exp\left[r\beta(1-k)(\overline{e} - \underline{e})\right], \tag{27}$$

where $\frac{c_2}{c_1} \equiv \frac{\partial c/\partial \underline{e}}{\partial c/\partial \overline{e}} = \frac{s\lambda\overline{e}+\lambda^2\underline{e}}{\overline{e}+s\lambda\underline{e}}$. Note that (27) is a generalized version of (4) in which the constant λ is replaced by the function $\frac{\partial c/\partial \underline{e}}{\partial c/\partial \overline{e}}$. Optimal efforts under EPD satisfy (26) and equation (9), with the left-hand side replaced by the left-hand side of (27).

Consider now the three environments which we studied in detail in Section 6. As $\lambda \to 1$ or as $r \to \infty$, $\sigma^2 \to 0$, both EAR and EPD induce perfectly balanced efforts for any $k \in (-1, \frac{c_1}{c_2})$, so the only effect of increasing k is to improve the diversification of risk from the shocks. Hence it is optimal in both environments to set k as large as possible subject to keeping efforts perfectly balanced: As $\lambda \to 1$, the optimal k approaches $\frac{c_1}{c_2} \to 1$, while as $r \to \infty$, $\sigma^2 \to 0$, the optimal k approaches $\frac{c_1}{c_2} \to \frac{1+s\lambda}{\lambda(\lambda+s)}$. Observe that in both cases, therefore, the optimal k approaches $1/v^{NHI}$. Therefore, just as in the original model, in these two limiting environments, optimally designed EAR and EPD generate a payoff for the principal arbitrarily close to what she achieves in the NHI benchmark. In the setting where the correlation of the shocks approaches 1, the choice of weight k has no effect on diversifaction, so it is optimal under EAR and EPD to set k to induce perfectly balanced efforts, and in this setting, too, optimally designed EAR and EPD generate a payoff arbitrarily close to that in the NHI benchmark.

In these limiting environments in which optimally designed ambiguous schemes induce perfectly balanced efforts, it follows that the ambiguous schemes dominate the best deterministic scheme. As long as $s \geq \frac{1}{\lambda}$, we saw above that under hidden information, no linear deterministic incentive scheme could induce both types of agent to choose relatively balanced efforts, and as a consequence, the payoff from the best deterministic scheme is bounded away from the payoff in the NHI benchmark. Since the ambiguous schemes generate a payoff arbitrarily close to the NHI benchmark, the ambiguous schemes perform better. Hence, allowing the agent's efforts on the tasks to be less than perfect substitutes in his cost function does not alter our main results.

7.3 Beyond The Exponential-Normal Model

Our findings that ambiguous incentive schemes induce more balanced efforts than symmetric deterministic ones and do so in a way that is more robust to hidden information on the agent's part apply even outside the exponential-normal model we have been considering, as we show in supporting online material.

8 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 ambiguous contracts can, in certain circumstances, alleviate this gaming. In such circumstances, ambiguity in the incentive scheme 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 ambiguous contracts dominate all deterministic incentive schemes. We identified three such environments. Each of these environments has the feature that optimally designed ambiguous 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 an ambiguous one. Furthermore, in these settings, optimally designed ambiguous schemes allow the principal to achieve a payoff arbitrarily close to what she could achieve in the absence of hidden information on the agent's part. That is, ambiguous schemes eliminate the efficiency losses from the agent's better knowledge of the environment.

It is worth noting that the outcomes achieved under ex ante randomization and ex post discretion in our model are achievable even if the principal cannot commit to a randomizing procedure in advance. The outcome under ex ante randomization is equivalent to the equilibrium outcome of a simultaneous-move game between the principal and the agent, and ex post discretion allows the principal to wait before choosing which performance measure to reward more highly until after she has observed outputs. Therefore, our ambiguous schemes are feasible even when the principal is unable to commit to complicated non-linear contracts. We suggest that part of the appeal of ambiguous contracts is that they approximate the outcomes of 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 ambiguous incentive schemes in more general environments.

References

[1] **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.

[2] **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.

[3] **Bentham, Jeremy**, *Constitutional Code*, Vol. 1, London: R. Heward, 1830.

[4] **Bernheim, B. Douglas and Michael D. Whinston**, "Incomplete Contracts and Strategic Ambiguity," *American Economic Review*, 1998, *88*, 902-932.

[5] **Bevan, Gwyn and Christopher Hood**, "Targets, Inspections, and Transparency," *British Medical Journal*, March 2004, *328*, 598.

[6] **Bull, Clive**, "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, Feb. 1987, *102* (1), 147-159.

[7] **Cain, Michael**, "The Moment-Generating Function of the Minimum of Bivariate Normal Random Variables," *The American Statistician*, May 1994, *48* (2), 124-125.

[8] **Gjesdal, Froystein**, "Information and Incentives: The Agency Information Problem," *Review of Economic Studies*, Jul. 1982, *49* (3), 373-390.

[9] **Grossman, Sanford J. and Oliver D. Hart**, "An Analysis of the Principal-Agent Problem," *Econometrica*, Jan. 1983, *51* (1), 7-45.

[10] **Holmström, Bengt**, "Moral Hazard and Observability," *Bell Journal of Economics*, Spring 1979, *10* (1), 74-91.

[11] —, "Moral Hazard in Teams," *Bell Journal of Economics*, Autumn 1982, *13* (2), 324-40.

[12] **— and Paul Milgrom**, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, Mar. 1987, *55* (2), 303-328.

[13] **— 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.

[14] **Jehiel, Philippe**, "On Transparency in Organizations, " *Working Paper*, 2011.

[15] **Jehiel, Philippe and Andrew F. Newman**, "Loopholes: Social Learning and the Evolution of Contract Form, " *Working Paper*, 2009.

[16] **Larkin, Ian**, "The Cost of High-powered Incentives: Salesperson Gaming in Enterprise Software," *HBS Working Paper*, 2006.

[17] **Lazear, Edward P.**, "Speeding, Terrorism, and Teaching to the Test," *Quarterly Journal of Economics*, Aug. 2006, *121* (3), 1029-1061.

[18] **Levin, Jonathan**, "Relational Incentive Contracts," *American Economic Review*, Jun. 2003, *93* (3), 835-857.

[19] **MacDonald, Glenn and Leslie M. Marx**, "Adverse Specialization," *Journal of Political Economy*, Aug. 2001, *109* (4), 864-899.

[20] **MacLeod, W. Bentley and James M. Malcomson**, "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, Mar. 1989, *57* (2), 447-480.

[21] **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.

[22] **Osler, M.,** "The Case Against Transparency in Law School Rankings Methodology," *Law School Innovation*, May 4, 2010.

[23] **Oyer, Paul**, "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality," *Quarterly Journal of Economics*, Feb. 1998, *113* (1), 149-185.

[24] **Prendergast, Canice**, "The Provision of Incentives in Firms," *Journal of Economic Literature*, Mar. 1999, 37 (1), 7-63.

[25] **Scott, Robert E. and George G. Triantis**, "Anticipating Litigation in Contract Design," *Yale Law Journal*, Jan. 2006, *115* (4), 814-879.

[26] **Weisbach, David A.**, "An Efficiency Analysis of Line Drawing in the Tax Law," *Journal of Legal Studies*, Jan. 2000, *29* (1), 71-97.

A Omitted Proofs

Proof of Lemma 1. The agent's certainty equivalent under the contract is

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

Given that the principal sets α to satisfy the agent's participation constraint with equality, the principal's expected payoff as a function of β is

$$\Pi^{SD}(\beta; \lambda = 1) = \frac{\beta}{2}\left(1 + \frac{1}{\delta}\right) - \frac{\beta^2}{2} - r\sigma^2\beta^2(1+\rho). \tag{28}$$

With β chosen optimally, the resulting maximized payoff is

$$\Pi^{SD}(\lambda = 1) = \frac{(\delta+1)^2}{8\delta^2\left[1 + 2r\sigma^2(1+\rho)\right]}. \tag{29}$$

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, her expected payoff as a function of β is

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

and the optimal choice of β yields payoff

$$\Pi^{OT} = \frac{1}{2\delta^2\left[1 + r\sigma^2\left(1 - \rho^2\right)\right]}. \tag{30}$$

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, $r\sigma^2\beta^2(1+\rho)$, is larger than that under the OT contract, $\frac{1}{2}r\sigma^2\beta^2\left(1-\rho^2\right)$. Therefore the principal faces a tradeoff between the more balanced efforts induced by SD and the lower risk imposed by OT. Comparison of (29) and (30) shows that there is a critical value of the principal's complementarity parameter δ,

$$\delta^1(r\sigma^2, \rho) \equiv 2\left[\frac{1 + 2r\sigma^2(1+\rho)}{1 + r\sigma^2(1-\rho^2)}\right]^{\frac{1}{2}} - 1, \tag{31}$$

above which the SD contract is optimal and below which the OT contract is preferred. The critical value

$\delta^1\left(r\sigma^2,\rho\right)$ is greater than 1 and is increasing in each of its arguments. ■

Proof of Lemma 2. Each agent type's certainty equivalent from contract C_i^{bal} is

$$ACE_i(C_i^{bal}) = E(w_i) - c_i(e_1,e_2) - \frac{1}{2}r\sigma^2 Var(w_i) = \alpha + \beta^2 - \frac{\beta^2}{2} - \frac{1}{2}r\sigma^2\beta^2(1+2\rho\lambda+\lambda^2).$$

The principal will set α to satisfy both types' participation constraint with equality, and her expected payoff from both types, as a function of β, will be

$$\Pi^{bal}(\beta) = \frac{\beta}{1+\lambda}\left(1+\frac{1}{\delta}\right) - \frac{\beta^2}{2} - \frac{1}{2}r\sigma^2\beta^2(1+2\rho\lambda+\lambda^2). \tag{32}$$

With β chosen optimally, the resulting maximized payoff is

$$\Pi^{bal} = \frac{(\delta+1)^2}{2\delta^2(1+\lambda)^2\left[1+r\sigma^2(1+2\rho\lambda+\lambda^2)\right]}. \tag{33}$$

This payoff is continuous as $\lambda \to 1$ and its limiting value is the payoff from the SD contract when λ equals 1.

The second type of contract pair which can be optimal in the "no hidden information benchmark" setting is a mirror-image pair of "one task" (OT) contracts:

$$\begin{aligned} C_1^{OT} &: \; w_1 = \alpha + \beta x_1 - \rho\beta x_2, \\ C_2^{OT} &: \; w_2 = \alpha + \beta x_2 - \rho\beta x_1. \end{aligned}$$

When assigned contract C_i^{OT}, each agent i chooses $e_i = \beta$ and $e_j = 0$. The principal's payoff will be the same from each type of agent, and with α and β chosen optimally, this payoff will be given by the expression (30) derived previously for the OT contract. Note that this payoff from the contract pair (C_1^{OT}, C_2^{OT}) is independent of λ, since neither type of agent exerts any effort on his more costly task.

Comparison of the payoff expressions (33) and (30) shows that there is a critical value of the principal's complementarity parameter δ,

$$\delta^{NHI}(\lambda, r\sigma^2,\rho) \equiv (\lambda+1)\left[\frac{1+r\sigma^2(1+2\rho\lambda+\lambda^2)}{1+r\sigma^2(1-\rho^2)}\right]^{\frac{1}{2}} - 1, \tag{34}$$

above which it is optimal to induce perfectly balanced efforts with (C_1^{bal}, C_2^{bal}) and below which it is optimal to induce fully focused efforts with (C_1^{OT}, C_2^{OT}). This critical value $\delta^{NHI}(\lambda, r\sigma^2, \rho)$ is increasing in each of its arguments and, as λ approaches 1, δ^{NHI} approaches δ^1. ■

Proof of Proposition 1. The certainty-equivalents that each of the two ADM contracts offers to each of the two types of agent are:

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

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

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

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

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

Since for all $\lambda > 1$ we have $ACE_1(C_2) \geq ACE_2(C_2)$, the second and fourth constraints above imply that the third constraint will not bind, and hence agent 1 earns an "information rent".[19]

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

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

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

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

$$
\text{s.t. } \beta_1 \geq \lambda\beta_2.
$$

It is not difficult to show that the constraint $\beta_1 \geq \lambda\beta_2$ will be binding at the optimum if and only if the principal's complementarity parameter δ is greater than or equal to $\hat{\delta}$, where

$$
\hat{\delta} \equiv \left(\frac{1+\lambda}{\lambda}\right)\left(\frac{\lambda^2 + r\sigma^2(\lambda^2+2\rho\lambda+1)}{1+r\sigma^2(1-\rho^2)}\right) - 1. \tag{35}
$$

If $\delta < \hat{\delta}$, then the principal's maximized payoff from this "unconstrained" asymmetric deterministic menu (ADMU) is

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

whereas if $\delta \geq \hat{\delta}$, then her maximized payoff from the "constrained" ADM (ADMC) is

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

If the principal wants to induce both types of agent to focus their effort on their preferred task, it is optimal to offer a menu of the following form, which we term a "Symmetric Deterministic Menu" or SDM:

$$
\begin{aligned}
C_1 &: w_1 = \alpha + \beta x_1 - \rho\beta x_2, \\
C_2 &: w_2 = \alpha + \beta x_2 - \rho\beta x_1.
\end{aligned}
$$

Faced with such a menu, agent i strictly prefers contract C_i to contract C_j and, having chosen C_i, will then set $e_i = \beta$ and $e_j = 0$. The negative coefficient in each contract on the agent's less-preferred task improves

[19] If the ADM were designed to induce balanced efforts from (only) agent 1, then agent 2 would be the one to earn an information rent.

insurance by exploiting the correlation in the shocks.

Clearly, this symmetric deterministic menu generates the same outcome, for each type of agent, as the principal achieves, in the no hidden information benchmark setting, from the contract pair (C_1^{OT}, C_2^{OT}). Thus, if the principal wants both types to exert fully focused effort on their preferred task, there is no loss to the principal from not knowing which task the agent actually prefers. Therefore, from equation (30), we know that the principal's maximized payoff from a SDM is

$$\Pi^{SDM} = \frac{1}{2\delta^2\left(1 + r\sigma^2\left(1 - \rho^2\right)\right)}, \tag{36}$$

which is independent of λ.

If, instead of using a symmetric deterministic menu, the principal were to use a SD contract, $w = \alpha + \beta x_1 + \beta x_2$, this would, for all $\lambda > 1$, also induce fully focused efforts from both types of agent but would impose a larger risk premium and hence generate a lower payoff for the principal. Maximized profit from a SD contract when $\lambda > 1$ would be

$$\Pi^{SD}(\lambda > 1) = \frac{1}{2\delta^2\left(1 + 2r\sigma^2\left(1 + \rho\right)\right)} \leq \Pi^{SDM} \qquad \forall\lambda > 1, \forall\delta. \tag{37}$$

Comparing the payoff expressions in (29) and (37) shows that the payoff from a SD contract drops discontinuously as the preference parameter λ is increased from 1, because the agent's efforts switch discontinuously from perfectly balanced to fully focused.

We now determine whether ADM or SDM are optimal. The comparison between SDM and ADMU involves comparing Π^{SDM} and Π^{ADMU}, where the former dominates if and only if

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

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

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

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

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

and since $r\sigma^2 \geq 0$, 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

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

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

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

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

The first-order conditions are

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

where

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

Adding the first-order conditions gives

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

$\Leftrightarrow$

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

so that we have

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

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

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

$\Leftrightarrow$

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

Solving the system of equations, we find for agent 1

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

For agent 2, analogous steps yield optimal effort levels

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

The maximized expected utilities are

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

where

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

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

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

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

$\Leftrightarrow$

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

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

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

where

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

The principal's expected payoff 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

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

In this expression for the principal's payoff, we have assumed that $\frac{1}{2} > p > \frac{1}{\lambda+1}$. Her payoff 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 payoff is increasing in p for $p \in (\frac{1}{1+\lambda}, \frac{1}{2})$, 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 $\overline{e}$ and the latter by $\underline{e}$, we have

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

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

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

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

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 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

$$\begin{aligned} 0 = & -r(e_1+\lambda e_2)m(-r\beta) \\ & +r\beta\exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right)\Phi\left(\frac{e_2-e_1+r\beta\sigma^2(1-\rho)}{\theta}\right) \\ & +\frac{1}{\theta}\exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right)\phi\left(\frac{e_2-e_1+r\beta\sigma^2(1-\rho)}{\theta}\right) \\ & -\frac{1}{\theta}\exp\left(-r\beta e_2 + \frac{1}{2}r^2\beta^2\sigma^2\right)\phi\left(\frac{e_1-e_2+r\beta\sigma^2(1-\rho)}{\theta}\right). \end{aligned} \quad (38)$$

Similarly for e_2 we have

$$\begin{aligned} 0 = & -\lambda r(e_1+\lambda e_2)m(-r\beta) \\ & +r\beta\exp\left(-r\beta e_2 + \frac{1}{2}r^2\beta^2\sigma^2\right)\Phi\left(\frac{e_1-e_2+r\beta\sigma^2(1-\rho)}{\theta}\right) \\ & +\frac{1}{\theta}\exp\left(-r\beta e_2 + \frac{1}{2}r^2\beta^2\sigma^2\right)\phi\left(\frac{e_1-e_2+r\beta\sigma^2(1-\rho)}{\theta}\right) \\ & -\frac{1}{\theta}\exp\left(-r\beta e_1 + \frac{1}{2}r^2\beta^2\sigma^2\right)\phi\left(\frac{e_2-e_1+r\beta\sigma^2(1-\rho)}{\theta}\right). \end{aligned} \quad (39)$$

Adding the two first order conditions we find

$$e_1 + \lambda e_2 = \frac{\beta}{(\lambda+1)}. \tag{40}$$

Expanding the third and fourth terms in the two first-order conditions (38) and (39) reveals that in both FOC's these terms net to 0 for all (e_1, e_2), and hence for (38) 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 (40) yields

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

$\Leftrightarrow$

$$\lambda = \exp\left[r\beta(e_1 - e_2)\right] \frac{\Phi\left(\frac{e_1 - e_2 + r\beta\sigma^2(1-\rho)}{\theta}\right)}{\Phi\left(\frac{e_2 - e_1 + r\beta\sigma^2(1-\rho)}{\theta}\right)}. \tag{41}$$

Both factors on the RHS of (41) 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 (41) 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 EPD 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 $\overline{e}^{EPD}$ and on his less-preferred task by $\underline{e}^{EPD}$. Define $d^{EPD} \equiv \overline{e}^{EPD} - \underline{e}^{EPD}$. Using (40) and (41), we can express the maximized expected utility of both types under ex post discretion as

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

$\Leftrightarrow$

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

For both types of agent, the certainty equivalent is

$$ACE = \alpha + \beta\underline{e}^{EPD} - \frac{\beta^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2\beta^2 - \frac{1}{r}\ln\left[\frac{1+\lambda}{\lambda}\Phi\left(\frac{\overline{e}^{EPD} - \underline{e}^{EPD} + r\beta\sigma^2(1-\rho)}{\theta}\right)\right] \tag{42}$$

while the principal's expected payoff is

$$
\begin{aligned}
\Pi^{EPD} &= \underline{e}^{EPD} + \frac{1}{\delta}\overline{e}^{EPD} - \alpha - \beta E\min\{x_1, x_2\} \\
&= \underline{e}^{EPD} + \frac{1}{\delta}\overline{e}^{EPD} - \alpha - \beta\left[\overline{e}^{EPD}\Phi\left(-\frac{d^{EPD}}{\theta}\right) + \underline{e}^{EPD}\Phi\left(\frac{d^{EPD}}{\theta}\right) - \theta\phi\left(\frac{d^{EPD}}{\theta}\right)\right] \\
&= \underline{e}^{EPD} + \frac{1}{\delta}\overline{e}^{EPD} - \alpha - \beta\underline{e}^{EPD} - \beta d^{EPD}\Phi\left(-\frac{d^{EPD}}{\theta}\right) + \beta\theta\phi\left(\frac{d^{EPD}}{\theta}\right).
\end{aligned} \tag{43}
$$

Using (42) to substitute into (43) yields the principal's expected payoff

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

■

Proof of Proposition 5. Equations (5) and (10) give the principal's payoff from interior effort choices by the agents under EAR and EPD, respectively, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$. The proof proceeds in three steps:

Step 1:

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

This inequality reflects the fact that for any given β and k, EPD 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 (10) is non-negative. Define $d \equiv (1-k)(\overline{e} - \underline{e})$. Then the sum in question has the sign of

$$
-r\beta d\Phi\left(\frac{-d}{\theta^k}\right) + r\beta\theta^k\phi\left(\frac{d}{\theta^k}\right) - \ln\left[\exp\{-r\beta d\}\Phi\left(\frac{-d}{\theta^k} + \frac{r\beta\theta^k}{2}\right) + \Phi\left(\frac{d}{\theta^k} + \frac{r\beta\theta^k}{2}\right)\right]. \tag{44}
$$

Now define $t \equiv \frac{r\beta\theta^k}{2}$ and $y \equiv \frac{d}{\theta^g}$. The expression (44) can be rewritten as

$$
h(y,t) \equiv -2ty\Phi(-y) + 2t\phi(-y) - \ln\left[\exp\{-2ty\}\Phi(-y+t) + \Phi(y+t)\right]. \tag{45}
$$

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 $\delta \geq \lambda$,

$$
\begin{aligned}
\underline{e}^{EPD} &+ \frac{1}{\delta}\overline{e}^{EPD} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^k)^2\beta^2 \\
&\geq \underline{e}^{EAR} + \frac{1}{\delta}\overline{e}^{EAR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^k)^2\beta^2
\end{aligned}
$$

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

Step 3: Note that

$$\begin{aligned}
\underline{e}^{EAR} &+ \frac{1}{\delta}\overline{e}^{EAR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^k)^2\beta^2 \\
&\geq \underline{e}^{EAR} + \frac{1}{\delta}\overline{e}^{EAR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma^k)^2\beta^2 \\
&\quad -\frac{1}{2r}\ln\left(\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right) \\
&= \Pi^{EAR}(\beta,k)
\end{aligned}$$

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, EAR 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 6.

Proof of Part 1: For $\lambda = 1$, both EAR and EPD induce interior solutions for efforts for all $\beta > 0$ and $k \in (-1, 1)$. Therefore, from Proposition 5, we know that EPD is more profitable than EAR for any given (β, k), so it suffices to show that, for any given (β, k), EPD can be dominated in terms of payoffs by a suitably designed symmetric deterministic (SD) scheme.

For $\lambda = 1$, aggregate effort under EPD is $\overline{e}^{EPD} + \lambda\underline{e}^{EPD} = \frac{\beta(1+k)}{2}$, and $\overline{e}^{EPD} = \underline{e}^{EPD} = \frac{\beta(1+k)}{4}$. Hence, for $\lambda = 1$, we can use equation (10) to write

$$\begin{aligned}
\Pi^{EPD}(\beta,k) = \left(\frac{\delta+1}{\delta}\right)\frac{\beta(1+k)}{4} &- \frac{1}{8}\beta^2(1+k)^2 - \frac{1}{2}r(\sigma^k)^2\beta^2 \\
&- \frac{1}{r}\left(\ln\left[2\Phi\left(\frac{r\beta\theta^k}{2}\right)\right] - r\beta\theta^k\phi(0)\right).
\end{aligned} \tag{46}$$

Consider now a SD scheme with incentive coefficient β^{SD} chosen to induce the same level of aggregate effort as under EPD 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 EPD. The principal's payoff under the SD scheme is

$$\begin{aligned}
\Pi^{SD}(\beta^{SD}) &= \frac{\delta+1}{\delta}\frac{\beta^{SD}}{2} - \frac{1}{2}\left(\beta^{SD}\right)^2 - r\sigma^2\left(\beta^{SD}\right)^2(1+\rho) \\
&= \frac{\delta+1}{\delta}\frac{\beta(1+k)}{4} - \frac{1}{8}\beta^2(1+k)^2 - \frac{1}{4}r\sigma^2\beta^2(1+k)^2(1+\rho)
\end{aligned} \tag{47}$$

Using equations (46) and (47) and the definitions of σ^k and θ^k in the statement of Proposition 4, we can write the difference in payoffs between the SD scheme and the EPD scheme as

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

Now, as in the proof of Proposition 5, define $t \equiv \frac{r\beta\theta^k}{2}$. Then

$$t^2 = \frac{r^2\sigma^2\beta^2(1-\rho)(1-k)^2}{2}, \tag{48}$$

and the payoff difference given by equation (48) has the sign of

$$g(t) \equiv \frac{t^2}{2} + \ln\left[2\Phi(t)\right] - 2t\phi(0). \tag{49}$$

Analyzing this function we have

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

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 $\rho < 1$, $\Pi^{SD}(\beta^{SD}) - \Pi^{EPD}(\beta, k) > 0$. If $\rho = 1$, then $\theta^k = 0$ for all $k \in (-1, 1)$, so $t = 0$, hence $\Pi^{SD}(\beta^{SD}) - \Pi^{EPD}(\beta, k) = 0$.

Proof of Part 2: Since Proposition 5 provides a payoff comparison between EAR and EPD only for the case where both schemes induce interior solutions for efforts, we analyze EAR and EPD separately to prove the assertions in Part 2.

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

$$\frac{\beta - \bar{e}^{EAR}}{\bar{e}^{EAR} - k\beta} = \exp\left\{r\beta\bar{e}^{EAR}(1-k)\right\}. \tag{50}$$

Since the RHS of (50) is > 1 for $k < 1$, (50) implies that

$$\bar{e}^{EAR} < \frac{\beta(1+k)}{2}. \tag{51}$$

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

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

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

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

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

Therefore

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

where the first inequality holds since $Z>1$ and the second follows from inequality (51).

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

$$\frac{\beta-\bar{e}^{EPD}}{\bar{e}^{EPD}-k\beta}=\exp\left\{r\beta(1-k)\bar{e}^{EPD}\right\}\frac{\Phi\left(\frac{(1-k)\bar{e}^{EPD}+r(\sigma^{k})^{2}\beta\left(1-\rho^{k}\right)}{\theta^{g}}\right)}{\Phi\left(\frac{-(1-k)\bar{e}^{EPD}+r(\sigma^{k})^{2}\beta(1-\rho^{k})}{\theta^{k}}\right)}. \tag{52}$$

Since the RHS of (52) is ≥ 1 for $k<1$, (52) implies that $\bar{e}^{EPD}\leq\frac{\beta(1+k)}{2}$. When EPD induces A to choose the corner solution $\left(\bar{e}^{EPD},0\right)$,

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

where

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

Consider now a SD scheme with incentive coefficient β^{SD} chosen to induce the same effort pair $\left(\bar{e}^{EPD},0\right)$

as under EPD for the given values of β and k: $\beta^{SD} = \bar{e}^{EPD}$. The principal's payoff under this SD scheme is

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

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

$$\begin{aligned}\frac{r^2\sigma^2}{4}[2\beta^2(1 &+ 2\rho k + k^2) - 4(1+\rho)\left(\bar{e}^{EPD}\right)^2] \\ &+ \ln\left[\Phi(+) + \exp\left\{-r\beta(1-k)\bar{e}^{EPD}\right\}\Phi(-)\right] \\ &+ r\beta(1-k)\bar{e}^{EPD}\Phi\left(\frac{-(1-k)\bar{e}^{EPD}}{\theta^g}\right) + r\beta\theta^g\phi\left(\frac{(1-k)\bar{e}^{EPD}}{\theta^g}\right). \qquad (53)\end{aligned}$$

Since (52) implies that $\bar{e}^{EPD} \leq \frac{\beta(1+k)}{2}$, the expression on the first line of (53) is greater than or equal to

$$\frac{r^2\sigma^2}{4}\left[2\beta^2(1+2\rho k+k^2) - (1+\rho)(1+k)^2\right] \qquad (54)$$

Now define $y \equiv \frac{(1-k)\bar{e}^{EPD}}{\theta^k}$ and $t \equiv \frac{r\beta\theta^k}{2}$. Then, using (54) and the second and third lines of (53), we conclude that

$$\Pi^{SD}\left(\beta^{SD}\right) - \Pi^{EPD}(\beta, k) \geq \frac{t^2}{2} - h(y,t),$$

where, as in the proof of Proposition 5,

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

In the proof of Proposition 5 it was noted that for all $y \geq 0$, $t \geq 0$, $h(y,t)$ is decreasing in y, and hence

$$\begin{aligned}\frac{t^2}{2} - h(y,t) &\geq \frac{t^2}{2} - h(0,t) \\ &= \frac{t^2}{2} + \ln\left[2\Phi(t)\right] - 2t\phi(0) \\ &= g(t),\end{aligned}$$

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 $t > 0$. Therefore, $\Pi^{SD}\left(\beta^{SD}\right) - \Pi^{EPD}(\beta, k) \geq 0$.

Proof of Part 3: There are two cases to consider: (i) EAR and EPD induce interior solutions for efforts or (ii) EAR and EPD induce corner solutions. The proof of Part 2 has dealt with the latter case, so here we treat the former. From Proposition 5, we know that EPD is more profitable than EAR 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), EPD can be dominated in terms of payoffs by a suitably designed symmetric deterministic (SD) scheme.

Define $d \equiv (1-k)(\bar{e}^{EPD} - \underline{e}^{EPD})$, $y \equiv \frac{d}{\theta^k}$, and $t \equiv \frac{r\beta\theta^k}{2}$. Then from the proof of Proposition 5, we know that we can write

$$\begin{aligned}
\Pi^{EPD}(\beta,k) &= \underline{e}+\frac{\bar{e}}{\delta}-\frac{1}{2}\left(\bar{e}+\lambda\underline{e}\right)^2-\frac{1}{2}r\beta^2(\sigma^k)^2+\frac{1}{r}h(y,t)\\
&\leq \underline{e}+\frac{\bar{e}}{\delta}-\frac{1}{2}\left(\bar{e}+\lambda\underline{e}\right)^2-\frac{1}{2}r\beta^2(\sigma^k)^2+\frac{1}{r}h(0,t)\\
&= \underline{e}+\frac{\bar{e}}{\delta}-\frac{1}{2}\left(\bar{e}+\lambda\underline{e}\right)^2-\frac{1}{2}r\beta^2(\sigma^k)^2+\frac{1}{2}[-\ln(2\Phi(t))+2t\phi(0)]\\
&\leq \frac{\bar{e}+\lambda\underline{e}}{\delta}-\frac{1}{2}\left(\bar{e}+\lambda\underline{e}\right)^2-\frac{1}{2}r\beta^2(\sigma^k)^2+\frac{1}{r}\left[-\ln\left(2\Phi\left(t\right)\right)+2t\phi\left(0\right)\right]\\
&= \frac{\beta(1+k)}{\delta\left(1+\lambda\right)}-\frac{1}{2}\frac{\beta^2}{\left(1+\lambda\right)^2}-\frac{1}{2}r\beta^2(\sigma^k)^2+\frac{1}{r}\left[-\ln\left(2\Phi\left(t\right)\right)+2t\phi\left(0\right)\right], \qquad (55)
\end{aligned}$$

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 EPD for the given values of β and k: $\beta^{SD}=\frac{\beta(1+k)}{1+\lambda}$. The principal's payoff under this SD scheme is

$$\Pi^{SD}(\beta^{SD})=\frac{\beta(1+k)}{\delta\left(1+\lambda\right)}-\frac{1}{2}\frac{\beta^2(1+k)^2}{\left(1+\lambda\right)^2}-r\sigma^2\beta^2(1+k)^2\frac{1+\rho}{(1+\lambda)^2} \qquad (56)$$

Hence from (55) and (56) we can conclude that

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

where the first inequality is a consequence of the inequalities in (55), the second inequality follows from the fact that $\lambda \geq 1$, the second equality uses (48), and the final line uses the definition of $g(t)$ in (49) and its non-nonegativity, as was established in the proof of Part 1 of this proposition. ■