

Gaming and Strategic Ambiguity in Incentive Provision

Florian Ederer*
UCLA

Richard Holden†
UNSW and NBER

Margaret Meyer‡
Oxford and CEPR

January 16, 2013

Abstract

It is often suggested that incentive schemes under moral hazard can be *gamed* by an agent with superior knowledge of the environment, and that deliberate lack of transparency about the incentive scheme can reduce gaming. We formally investigate these arguments. Ambiguous incentive schemes induce more balanced efforts from an agent who performs multiple tasks and is better informed about the environment, but also impose more risk on the agent. If tasks are sufficiently complementary for the principal, ambiguous 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*: exploitation of an incentive scheme by an agent for his own self-interest, to the detriment of the objectives of the incentive designer. Gaming can take numerous forms, among them i) diversion of effort away from activities that are socially valuable but difficult to measure and reward, towards activities that are easily measured and rewarded; ii) exploitation of the rules of classification to improve apparent, though not actual, performance; or iii) distortion of choices about timing and/or pricing to exploit temporarily high monetary rewards even when socially efficient choices have not changed. Evidence of the first type of gaming is provided by Burgess, Propper, Ratto, and Tominey (2012) and Carrell and West (2010), of the second type by Gravelle, Sutton, and Ma (2010), and of the third type by Oyer (1998) and Forbes, Lederman, and Tombe (2012).¹ The costs of gaming are exacerbated when the agent has superior knowledge of the environment: this makes the form and extent of gaming harder to predict and hence harder to deter.

It has been suggested that lack of transparency—being deliberately ambiguous about the criteria upon which rewards will be based—can help deter gaming. This idea has a long intellectual history. It dates back at least to Bentham (1830), who argued that deliberate ambiguity about the content of civil service

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

†University of New South Wales, Australian School of Business, Sydney, NSW 2052, Australia, richard.holden@unsw.edu.au.

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

¹Burgess et al (2012) and Gravelle et al (2010) study UK public sector organizations (an employment agency and the National Health Service, respectively), Carrell and West (2010) study postsecondary education, while Oyer (1998) and Forbes et al (2012) study private sector organizations (salespeople and executives across various industries, and airlines, respectively).

selection tests would lead to the “maximization of the inducement afforded to exertion on the part of learners, by impossibilizing the knowledge as to what part the field of exercise the trial will be applied to, and thence making aptitude of equal necessity in relation to every part”.²

More recently, in the light of apparent gaming of incentive schemes introduced by the UK government to give hospitals stronger incentives to reduce patient waiting times, Bevan and Hood, in an editorial in the *British Medical Journal*, have argued, “What is needed are ways of limiting gaming. And one way of doing so is to introduce more randomness in the assessment of performance, at the expense of transparency” (2004, p. 598). Relatedly, Dranove, Kessler, McClellan, and Satterthwaite (2003) document that in the US, report cards for hospitals “encourage providers to ‘game’ the system by avoiding sick patients or seeking healthy patients or both” (p. 556), and they argue that such gaming is facilitated by providers having better information about patients’ characteristics than do the analysts who compile the report cards.

There is also increasing concern about gaming, by law school deans, of the data used by *U.S. News* to produce its highly influential law school rankings. In an attempt to mitigate gaming, law scholars (e.g., Osler 2010) have argued against transparency in the rankings methodology, and *U.S. News* itself announced, “The problem with being totally transparent about key methodology details about U.S. News’s America’s Best Law Schools Rankings is that it’s clear [...] that certain law schools are taking advantage of that knowledge to game our rankings. *U.S. News* is going to take steps to prevent these data manipulations by law schools in future rankings.”³

Relatedly, Google has experienced manipulation of its search results by canny retailers. For example, JC Penney was found to engage in “black hat optimization, the dark art of raising the profile of a Web site with methods that Google considers tantamount to cheating”,⁴ while DecorMyEyes.com rose to the top of Google’s search rankings by generating publicity through intentional harassment of its own customers.^{5,6} Although many retailers have been demanding greater transparency from Google about the algorithm generating its search results, Google “has become less and less transparent regarding the data that it chooses to make available [...] for security/privacy reasons”.⁷

One view as to why courts often prefer standards—which are somewhat vague—to specific rules is that standards mitigate 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 contracts can improve parties’ incentives to fulfill the spirit of the contract rather than just satisfy the narrowly defined stipulations.

There are numerous other examples of the deliberate use of vagueness in related incentive provision settings. The use of speed cameras is often randomized, and security checks at airports and tax audits are often random. Randomization is also routinely used in economic experiments: to encourage subjects to concentrate throughout the whole experiment while also keeping the expected experimental expenditures

²Bentham, 1830/2005, Ch. IX, §16, Art 60.1.

³U.S. News, May 20, 2010, <http://www.usnews.com/education/blogs/college-rankings-blog/2010/05/20/us-news-takes-steps-to-stop-law-schools-from-manipulating-the-rankings>

⁴New York Times, February 12, 2011, <http://www.nytimes.com/2011/02/13/business/13search.html>

⁵New York Times, November 26, 2010, <http://www.nytimes.com/2010/11/28/business/28borker.html>

⁶Jehiel and Newman (2011) develop a dynamic model in which principals learn from agents’ behavior about the possibilities for gaming of incentive schemes.

⁷Structural Search Engine Optimization, Google Penalty Solutions, November 4, 2011, <http://www.re1y.com/blog/occupy-google-blog.html>

low, subjects are often paid based on their performance in one randomly chosen period of the experimental session.⁸ In international relations, the policy of strategic ambiguity is described as “the practice by a country of being intentionally ambiguous on certain aspects of its foreign policy [...]. It may be useful if the country has contrary foreign and domestic policy goals or if it wants to take advantage of risk aversion to abet a deterrence strategy.”^{9,10}

In this paper, we construct a formal theory of gaming by an agent with superior knowledge of the environment, and we identify circumstances in which ambiguity, or lack of transparency about the weighting scheme used to determine rewards, can be beneficial in mitigating gaming. In our model, the risk-averse 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’s (1991) multi-task principal-agent model. Crucially, in our model, unlike in Holmström and Milgrom’s, 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, he 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, we study the performance of two simple types of “ambiguous” incentive scheme. Such schemes confront the agent with ambiguity, in that he knows that the compensation schedule ultimately used will take one of two possible linear forms, but when choosing 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 the scheme we term *ex ante randomization*, the principal chooses randomly, before outputs are observed, which compensation schedule to use. Under the one we term *ex post discretion*, the principal chooses which schedule to use after observing outputs on the two tasks. The source of the ambiguity confronting the agent is different under these two schemes, but the nature and comparative statics of the incentive effects are revealingly similar.

Ex ante randomization pushes the risk-averse 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. Ex post discretion, too, provides the agent with a self-insurance motive but also provides an additional incentive for effort balance: the principal’s strategic ex post choice of which of the two compensation schedules to use 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. For both types of ambiguous incentive scheme, we show that the more unequal are the weights on the performance measures in

⁸Take the particular example of Healy and Pate (2011), who explain in their instructions to subjects, “In the experiment today you will be asked to complete four different tasks [...] We will randomly select one of the tasks and pay you based on your performance in that task. Once you have completed the four tasks, we determine which task counts for payment by drawing a number between 1 and 4.” (p. 1 of Technical Appendix)

⁹Wikipedia, “Policy of Deliberate Ambiguity”, http://en.wikipedia.org/wiki/Policy_of_deliberate_ambiguity. Multiple objectives of the principal and risk aversion of the agent will also be important features of our model.

¹⁰In the context of arms proliferation, Baliga and Sjöström (2008) show that a small country may be able to use strategic ambiguity about whether it possesses weapons of mass destruction as a substitute for actual investment and thereby help to deter an attack.

the two possible compensation schedules, the stronger are the agent’s incentives to choose balanced efforts.

We demonstrate that ambiguous contracts are more robust to uncertainty about the agent’s preferences than are deterministic ones. With ambiguous contracts, the efforts exerted on the two tasks vary continuously when we introduce a small amount of such uncertainty, whereas they vary 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. While ex post discretion imposes lower risk than ex ante randomization, still, for any level of *aggregate* effort to be induced on the two tasks, a deterministic contract does so with the imposition of lower risk costs than either type of ambiguous scheme. As a consequence, as we show in Proposition 6, if the principal faces no uncertainty about the agent’s preferences, then both types of ambiguous scheme can be dominated by a linear deterministic one. When, however, the agent’s preferences are private information, the principal faces a trade-off between the stronger incentives for balanced efforts under ambiguous schemes and the lower risk costs under deterministic ones.

Our key contribution is to identify settings in which both of our simple types of ambiguous incentive scheme dominate all deterministic ones. We identify three such environments. In each one, optimally weighted ambiguous contracts induce both types of agent to choose perfectly balanced efforts on the two tasks. In the first such setting, 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 agent’s risk aversion becomes infinitely large and the variance of the shocks to outputs becomes arbitrarily small. In the final setting, the shocks to outputs become perfectly correlated. In all three of these settings, as we show in Propositions 7, 8, and 9, both ex ante randomization and ex post discretion allow the principal to achieve a payoff arbitrarily close to what he could achieve in the absence of hidden information on the agent’s part. That is, in these settings, both of the simple types of ambiguous incentive scheme completely eliminate the efficiency losses from the agent’s better knowledge of the environment.

1.1 Related Literature

Our paper builds on the theoretical analyses of Holmström and Milgrom (1987, 1991). The first of these provides conditions in a dynamic moral hazard setting 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. Discussing Mirrlees’s (1974) result that the first-best outcome in a hidden-action model can be approximated by a step-function (hence highly non-linear) incentive scheme, they argue “to construct the [Mirrlees] scheme, the principal requires very precise knowledge about the agent’s preferences and beliefs, and about the technology he controls. The two-wage scheme performs ideally if the model’s assumptions are precisely met, but can be made to perform quite poorly if small deviations in the assumptions [...] are introduced” (p. 305). Motivated both by the simplicity of linear contracts and by Holmström and Milgrom’s robustness arguments, we focus our analysis on compensation schedules in which, ex post, after all choices are made and random variables are realized, payments are linear functions of the performance measures.

Multi-task principal-agent models (e.g., Holmström and Milgrom (1991), Baker (1992)) have highlighted

that precise incentive contracts based on verifiable performance measures can be distortionary. When efforts on different tasks are technological substitutes for the agent, incentives on one task crowd out incentives on others, and as a result, even on easily measured tasks, optimal incentives may be low-powered. Models of self-enforcing contracts (e.g., MacLeod and Malcolmson (1989), Baker, Gibbons, and Murphy (1994), Bernheim and Whinston (1998)) have shown how agency costs can be reduced by allowing the principal to respond in a discretionary fashion to indicators of the agent's behavior that are observable but non-contractible. In our model, the ambiguous incentive scheme that we term *ex post* discretion can be beneficial even when the performance measures are contractible. None of the models above studies the potential benefits of exogenous randomization, as is used in the scheme we term *ex ante* randomization.

In general single-task hidden-action models allowing arbitrarily complex contracts, Gjesdal (1982) and Grossman and Hart (1983) show that exogenous randomization may be optimal, but only if the agent's risk tolerance varies with the level of effort he exerts. In our model, the agent's risk tolerance is independent of his effort level, and the attractiveness of *ex ante* randomization and *ex post* discretion stems from their ability to mitigate the agency costs of *multi-task* incentive problems.

The potential benefits of exogenous randomization have also been explored in hidden-information models, especially those studying the design of optimal tax schedules. Stiglitz (1982) and Pestieau, Posen, and Slutsky (1997), among others, have shown that randomization can facilitate the screening of privately-informed individuals and is especially effective when private information is multi-dimensional. In our hidden-action cum hidden-information setting, in contrast, *ex ante* randomization in fact eliminates the need for screening, as we show in Section 7.3.

The costs and benefits of transparency in incentive design are also explored in Jehiel (2011), Rahman (2012), and Lazear (2006). Jehiel (2011) shows in an abstract moral hazard setup that a principal may gain by keeping agents uninformed about some aspects of the environment (e.g., how important specific tasks are). The benefits of suppressing information in relaxing incentive constraints can outweigh the costs of agents' less efficient adaptation of actions to the environment. Rahman (2012) examines how a principal can provide incentives for an individual tasked with monitoring and reporting on the behavior of an agent. He shows that randomization by the principal over what he asks the agent to do allows the principal to incentivize the monitor effectively. Finally, Lazear (2006), in a model in which agents have no hidden information, explores high-stakes testing in education and the deterrence of speeding and terrorism, identifying conditions under which a lack of transparency can have beneficial incentive effects. In the testing context, a crucial role is played by the exogenous restriction on the number of topics that can be tested, whereas in our model, even when all tasks can be measured and rewarded, deliberate ambiguity about the weights in the incentive scheme can be desirable.

The paper perhaps most closely related to ours is MacDonald and Marx (2001). Like us, they analyze a principal-agent model with two tasks where the agent's efforts on the tasks are substitutes for the agent but complements for the principal, and where the agent is privately informed about his preferences. Since task outcomes are binary, contracts are automatically linear in each outcome and specify at most four distinct payments. In this simple environment, it is possible to solve for the optimal contract, and they show that the more complementary the tasks are for the principal, the more the optimal reward scheme makes successes on the tasks complementary for the agent. They assume that the principal can commit to such a

nonseparable contract. In fact, their optimal outcome could be implemented using the scheme we term ex post discretion, which requires less commitment power and uses (ex post) payment schemes that are simpler, because they are separable in the task outcomes. Moreover, because of their restriction to binary outcomes, ex ante randomization over separable payment schemes would have no impact on agents’ incentives in their model, whereas our analysis readily reveals that even beyond the exponential-normal setting on which we focus, ex ante randomization will in general provide incentives for agents to balance their efforts.

Section 2 outlines the model, and Section 3 analyzes the optimal deterministic incentive scheme. Section 4 studies ex ante randomization and ex post discretion. Section 5 identifies settings in which ambiguous contracts are dominated by deterministic schemes. Section 6, which is the heart of the paper, identifies environments in which optimally weighted ambiguous contracts dominate the best deterministic scheme. Section 7 discusses extensions and robustness of our results, and Section 8 concludes. Proofs not provided in the text are in the appendix.

2 The Model

A principal hires an agent to perform a job for him. 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 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 $(\varepsilon_1, \varepsilon_2)$ have a symmetric bivariate normal distribution with mean 0, variance σ^2 , and covariance $\rho\sigma^2 \geq 0$. The efforts chosen by the agent are not observable by the principal. In addition, at the time of contracting, the agent is privately informed about his cost 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. The parameter λ is common knowledge, and $\lambda \geq 1$. For each type of agent $i = 1, 2$, efforts are perfect substitutes: $\frac{\partial c_i / \partial e_1}{\partial c_i / \partial e_2}$ does not vary with (e_1, e_2) .¹¹ 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 $U = -\exp\{-r(w - c_i(e_1, e_2))\}$, 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

$$\Pi = \min\{e_1, e_2\} + \frac{1}{\delta} \max\{e_1, e_2\} - w.$$

The larger is the parameter $\delta \geq 1$, the more complementary are the agent’s efforts on the tasks. In the extreme case where $\delta = \infty$, the principal’s payoff function reduces to $\Pi = \min\{e_1, e_2\} - w$, and the efforts are perfect complements. At the other extreme, when $\delta = 1$, $\Pi = e_1 + e_2 - w$, so the efforts are perfect

¹¹In Section 7.4, we show that our key results are robust to a relaxation of this assumption.

substitutes—here the principal is indifferent as to how the agent allocates his total effort across the tasks.¹²

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 in the socially efficient effort allocation, each type of agent focuses exclusively on his preferred task.

The principal’s benefit $\min\{e_1, e_2\} + \frac{1}{\delta} \max\{e_1, e_2\}$ is assumed non-verifiable. Therefore, the only measures on which the agent’s compensation can be based are x_1 and x_2 . The principal chooses a compensation scheme to maximize his expected payoff, subject to participation and incentive constraints for the agent that reflect the agent’s hidden information and hidden actions. Incentive schemes will be compared according to the (expected) payoff generated for the principal.

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 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 examine how the form of the optimal deterministic scheme depends on the parameters of the environment, specifically the values of λ (the strength of each type of agent’s preferences across tasks), δ (the strength of the principal’s preference for balanced efforts), ρ (the correlation of the shocks to performance measures), and $r\sigma^2$ (the importance of risk aversion).¹³

Section 4 introduces the two simple 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 EAR, 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.

¹²We assume throughout that difficulties of coordination would prevent the principal from splitting the job between two agents, with each agent responsible for only one dimension (task).

¹³For deterministic schemes, the values of r and σ^2 affect the principal’s payoff only through their product $r\sigma^2$, but we show that for ambiguous schemes, r and σ^2 have separate influences on effort choices and hence on the principal’s payoff.

3 Deterministic Contracts

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

When $\lambda = 1$, 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 case, 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 from the random shocks affecting both tasks. The second 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 $w = \alpha + \beta x_1 - \rho \beta x_2$.

The SD contract makes the agent indifferent over all effort pairs that equate the common marginal cost of effort on the two tasks, $e_1 + e_2$, to the common marginal benefit, β . The principal prefers the agent to choose $e_1 = e_2 = \frac{\beta}{2}$, since the parameter $\delta \geq 1$, and we assume that the agent does indeed choose this balanced effort allocation.

Lemma 1 *When $\lambda = 1$, there exists a critical $\delta^1(r\sigma^2, \rho)$, strictly greater than 1 and increasing in each of its arguments, such that for $\delta > \delta^1$, the principal’s payoff is maximized by an optimally designed SD contract, and for $\delta < \delta^1$, it is maximized by an optimally designed OT contract.*

The SD contract induces the agent to exert effort on both tasks, while the OT contract elicits effort on only 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, $\frac{1}{2}r\sigma^2\beta^2(1 - \rho^2)$. Therefore, the principal faces a trade-off between the more balanced efforts induced by SD and the lower risk imposed by OT. If and only if his complementarity parameter exceeds $\delta^1(r\sigma^2, \rho)$, the benefits of the balanced efforts elicited by the SD contract outweigh the costs of the extra risk imposed on the agent.

3.2 Two Types of Agent but No Hidden Information

Suppose now that $\lambda > 1$, so there are two distinct types of agent, but the principal can observe the agent’s type and offer each type a different contract. We will refer to this as the “no hidden information benchmark” (NHI). This benchmark is important because, as we will see, there are environments in which optimally designed ambiguous contracts allow the principal, even in the presence of hidden information, to achieve a payoff arbitrarily close to that achievable in this benchmark.

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 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:

$$C_1^{bal} : w_1 = \alpha + \beta x_1 + \lambda \beta x_2 \quad \text{and} \quad C_2^{bal} : w_2 = \alpha + \beta x_2 + \lambda \beta x_1.$$

Since C_i^{bal} , the contract assigned by the principal to the type- i agent, rewards type i ’s more costly task at a rate λ times as large as his less costly task, type 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 agent types the contract C_i^{bal} approaches the symmetric deterministic (SD) contract, and the efforts induced by C_i^{bal} as $\lambda \rightarrow 1$ approach those induced by SD when $\lambda = 1$.

The second type of contract pair which can be optimal in the “no hidden information benchmark” 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:

$$C_1^{OT} : w_1 = \alpha + \beta x_1 - \rho \beta x_2 \quad \text{and} \quad C_2^{OT} : w_2 = \alpha + \beta x_2 - \rho \beta x_1.$$

When assigned contract C_i^{OT} , each type i chooses $e_i = \beta$ and $e_j = 0$. The principal’s payoff from the contract pair (C_1^{OT}, C_2^{OT}) is independent of λ .

In choosing between (C_1^{bal}, C_2^{bal}) and (C_1^{OT}, C_2^{OT}) in the NHI setting, the principal faces a trade-off between the more balanced efforts induced by the former and the lower risk cost imposed by the latter.

Lemma 2 *In the no hidden information (NHI) benchmark, there exists a critical $\delta^{NHI}(\lambda, r\sigma^2, \rho)$, increasing in each of its arguments, such that for $\delta > \delta^{NHI}$, the principal’s payoff is maximized by an optimally designed pair of contracts of the form (C_1^{bal}, C_2^{bal}) , and for $\delta < \delta^{NHI}$, it is maximized by an optimally designed pair of contracts of the form (C_1^{OT}, C_2^{OT}) . As $\lambda \rightarrow 1$, $\delta^{NHI}(\lambda, r\sigma^2, \rho) \rightarrow \delta^1(r\sigma^2, \rho)$. Moreover, as $\lambda \rightarrow 1$, the principal’s payoff from optimally designed contracts of the form (C_1^{bal}, C_2^{bal}) approaches his maximized payoff under a symmetric deterministic (SD) contract.*

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 a screening device. In principle, three possible patterns of effort could emerge: (a) both types of agent exert balanced efforts, (b) one type exerts balanced efforts and the other type focused effort, or (c) both types exert focused effort.

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.*

The proof of Lemma 3 begins by noting that if such a menu existed, it would have the form

$$C_1 : w_1 = \alpha_1 + \beta_1 x_1 + \lambda \beta_1 x_2 \quad \text{and} \quad C_2 : w_2 = \alpha_2 + \beta_2 x_2 + \lambda \beta_2 x_1.$$

However, we show that no matter how $(\alpha_1, \beta_1, \alpha_2, \beta_2)$ were chosen, at least one type of agent would have an incentive to select the “wrong” contract from the menu and exert effort only on his preferred task. Lemma 3 implies that the only possible patterns of effort are (b) and (c) above, and hence that we can confine attention to schemes that are either (b) an “asymmetric deterministic menu” (ADM)

$$C_1^{ADM} : w_1 = \alpha_1 + \beta_1 x_1 - \rho \beta_1 x_2 \quad \text{and} \quad C_2^{ADM} : w_2 = \alpha_2 + \beta_2 x_2 + \lambda \beta_2 x_1,$$

or (c) a “symmetric deterministic menu” (SDM)

$$C_1^{SDM} : w_1 = \alpha + \beta x_1 - \rho \beta x_2 \quad \text{and} \quad C_2^{SDM} : w_2 = \alpha + \beta x_2 - \rho \beta x_1.$$

Under an SDM, each agent chooses the schedule with a positive coefficient on output on his preferred task and exerts effort only on that task; the coefficient for output on the other task is used for insurance. Under an ADM of the form above, one agent (here, type 2) is induced (through appropriate choice of (α_1, α_2)) to select schedule C_2^{ADM} , which leaves him willing to exert balanced effort on the two tasks; the type-1 agent selects C_1^{ADM} , which induces him to choose fully focused efforts (and uses the coefficient on x_2 to provide insurance). Relative to an SDM, an ADM has the benefit of inducing one type of agent to choose balanced efforts, but it imposes more risk on that agent type and also necessitates leaving rents to the other type. Whether the benefit of the ADM outweighs these costs depends on δ , the strength of the principal's preference for balanced efforts.

Proposition 1 *(i) When $\delta > \delta^{NHI}$, the principal is strictly worse off when hidden information is present than when it is absent. (ii) When the agent is privately informed about his preferences, there exists a critical $\delta^{HI}(\lambda, r\sigma^2, \rho)$, increasing in each of its arguments, such that for $\delta > \delta^{HI}$, the principal's payoff is maximized by an optimally designed ADM, and for $\delta < \delta^{HI}$, it is maximized by an optimally designed SDM. (iii) For all $\lambda > 1$ and for all $(r\sigma^2, \rho)$, $\delta^{HI} > \delta^{NHI}$, and as $\lambda \rightarrow 1$, $\delta^{HI}(\lambda, r\sigma^2, \rho) \rightarrow \delta^1(r\sigma^2, \rho)$. (iv) For $\delta > \delta^1(r\sigma^2, \rho)$, the limit as $\lambda \rightarrow 1$ of the principal's payoff from an optimally designed ADM is strictly below his maximized payoff under a SD contract.*

Part (i) of the proposition follows from the facts that, for $\delta > \delta^{NHI}$, it is a strict optimum for the principal in the NHI benchmark to induce both types of agent to choose perfectly balanced efforts and that this outcome is infeasible under hidden information. The result that $\delta^{HI}(\lambda, r\sigma^2, \rho) > \delta^{NHI}(\lambda, r\sigma^2, \rho)$, for all $\lambda > 1$, says that the principal's complementarity parameter δ must be larger when hidden information is present than when it is absent for him to find it optimal 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. Part (iv) shows that under hidden information, when $\delta > \delta^1(r\sigma^2, \rho)$, 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 privately-known preference across tasks, of inducing balanced efforts from both types with a deterministic scheme. In contrast, in the NHI benchmark, where it is feasible to induce balanced efforts from both types for all λ , Lemma 2 showed that the principal's maximized payoff falls continuously as λ is increased from 1.

4 Ambiguous Contracts

4.1 Ex Ante Randomization

A contract with *ex ante randomization* (EAR) involves two parameters, $\beta > 0$ and $k \in (-1, 1)$, and 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 $w = \alpha + \beta x_2 + k\beta x_1$. 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 tasks,

the principal's randomization exposes the agent to extra risk, risk against which he can insure himself by choosing more balanced efforts. By varying k , the principal can affect how much risk the randomization per se imposes on the agent and can thereby affect the strength of the agent's incentives to balance his efforts. If k were equal to 1, the randomized scheme would collapse to the symmetric deterministic (SD) contract 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 more balanced 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. When EAR induces interior solutions for efforts,

(ii) each type of agent chooses effort on his less costly task, \bar{e}^{EAR} , and effort on his more costly task, \underline{e}^{EAR} , satisfying

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

$$\exp[r\beta(1-k)(\bar{e}^{EAR} - \underline{e}^{EAR})] = \frac{\lambda-k}{1-k\lambda}; \quad (2)$$

(iii) the gap in efforts, $\bar{e}^{EAR} - \underline{e}^{EAR}$, is increasing in λ , approaching 0 as $\lambda \rightarrow 1$; decreasing in $r\beta$, approaching 0 as $r\beta \rightarrow \infty$; and increasing in k , approaching 0 as $k \rightarrow -1^+$;

(iv) the principal's payoff under EAR, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$, is

$$\Pi^{EAR}(\beta, k) = \underline{e}^{EAR} + \frac{1}{\delta} \bar{e}^{EAR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2} r \sigma^2 \beta^2 (1+2\rho k+k^2) - \frac{1}{2r} \ln \left[\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)} \right]. \quad (3)$$

To understand part (i), 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 agent types and for any effort levels. Hence in this case, both types of agent would optimally exert effort only on their preferred task.

To understand equation (1), 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 \bar{e} and \underline{e} must yield $\beta(1+k) = \partial c/\partial \bar{e} + \partial c/\partial \underline{e}$ for both types of agent. Since $\partial c/\partial \bar{e} + \partial c/\partial \underline{e} = (1+\lambda)(\bar{e} + \lambda \underline{e})$, this gives us (1). Throughout, we will refer to the quantity $\bar{e} + \lambda \underline{e}$ as an agent's *aggregate effort*, since it is the quantity which determines his total cost of effort.

Substituting (1) into either of the first-order conditions for optimal efforts yields

$$\frac{\lambda-k}{1-k\lambda} = \frac{E[U'(\cdot)I_{\{\underline{x} \text{ is rewarded more highly than } \bar{x}\}}]}{E[U'(\cdot)I_{\{\bar{x} \text{ is rewarded more highly than } \underline{x}\}}]}, \quad (4)$$

where \bar{x} (respectively, \underline{x}) denotes performance on an agent's less costly (respectively, more costly) task. Since under EAR, each compensation schedule is employed with probability one-half,

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

which when combined with (4) gives (2).

Equation (2) shows how the agent’s optimal degree of self-insurance against the compensation risk imposed on him by EAR varies with his preferences and with the parameters of the randomized contract. The smaller the cost difference between tasks (the smaller is λ), the less costly it is for the agent to self-insure by choosing relatively balanced efforts, so the smaller is his optimal gap between efforts on the two tasks. The more risk-averse the agent (the larger is r) or the larger the incentive intensity β , the more costly for the agent is the risk imposed by the randomization, so the stronger is his incentive to self-insure and the smaller his optimal effort gap. As either $\lambda \rightarrow 1$ or $r\beta \rightarrow \infty$, the optimal gap approaches 0, which corresponds to full self-insurance. Finally, the smaller is the parameter k , the more different are the two possible compensation schedules and the more costly is the risk imposed by the randomization, so the smaller is the optimal effort gap. As $k \rightarrow -1^+$, the self-insurance motive approaches its strongest level, and the optimal effort gap approaches 0.

Under a symmetric deterministic contract, the agent’s efforts change discontinuously as hidden information about the agent’s preferences is introduced, i.e., as λ is increased from 1: At $\lambda = 1$, efforts are perfectly balanced (the allocation preferred by the principal), but for any $\lambda > 1$, efforts are completely focused on a single task. As a result, the principal’s payoff from a SD contract drops discontinuously as λ is raised from 1. Furthermore, as Proposition 1 showed, even when the principal chooses a deterministic scheme optimally as a function of λ (as long as the optimal such scheme is an ADM), 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 efforts and the principal’s payoff are continuous in λ at $\lambda = 1$, as long as the agent is risk-averse. Thus EAR is more robust to the introduction of private information on the part of the agent than is the best deterministic scheme.¹⁴ 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 , his payoff will be discontinuously lower than if he were right. The performance of ex ante randomization does not display this extreme sensitivity.¹⁵

We have established Proposition 2 under the assumption that the principal can commit to randomizing uniformly between the two compensation schedules.¹⁶ 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 2.¹⁷ Thus the

¹⁴In Section 7.5 we show that, even outside the exponential-normal framework we have been using, EAR induces more balanced efforts than a SD contract and is more robust to the introduction of hidden information on the agent’s part.

¹⁵Bond and Gomes (2009) also study a multi-task principal-agent setting in which a small change in the agent’s preferences can result in a large change in the behavior induced by a contract and a consequent large drop in the principal’s payoff, a situation they term “contract fragility”.

¹⁶Given 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, assessed ex ante, and also avoids leaving any rent to either type of agent.

¹⁷To 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 profiles are mirror images, the principal anticipates equal expected output on the two tasks, so is willing to randomize uniformly over the two mirror-image compensation schedules. Given that the principal randomizes uniformly, the agents’ optimal behavior is clearly as described in the proposition. To see that this outcome is the *unique* equilibrium, observe that if the agents conjectured that the principal would choose the schedule rewarding task 1 more highly than task 2 with a probability greater than (less than) $1/2$, then their optimal efforts

attractive properties of ex ante randomization are 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 (3), the last two terms represent the total cost of the risk borne by the agent under EAR. The penultimate term 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 in (3). 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 (1) shows, aggregate effort $\bar{e} + \lambda \underline{e}$ is proportional to B . Using this definition and equations (1) and (2), we can re-express the principal's payoff (3) as a function of B and k :

$$\Pi^{EAR}(B, k) = \underline{e}^{EAR} + \frac{\bar{e}^{EAR}}{\delta} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2} r \sigma^2 B^2 \frac{1+2\rho k+k^2}{(1+k)^2} - \frac{1}{2r} \ln \left[\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)} \right], \quad (5)$$

where

$$\underline{e}^{EAR} + \frac{\bar{e}^{EAR}}{\delta} = \frac{\delta+1}{\delta} \frac{B}{(\lambda+1)^2} - \frac{\delta-\lambda}{\delta} \frac{\ln \left(\frac{\lambda-k}{1-k\lambda} \right)}{(\lambda+1) r B \left(\frac{1-k}{1+k} \right)}. \quad (6)$$

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. Equations (5) and (6) show that increasing k has three effects. First, a larger k raises the effort gap $\bar{e} - \underline{e}$ and, with B and hence aggregate effort $\bar{e} + \lambda \underline{e}$ held fixed, this larger gap lowers the principal's benefit $\underline{e} + \frac{\bar{e}}{\delta}$ whenever $\delta > \lambda$, i.e., whenever balanced efforts are socially efficient. 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. This second effect of k also reduces the principal's payoff and is reflected in the final term in (5). 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 $\Pi^{EAR}(B, k)$, as reflected in the second-to-last term in (5).

In general, the optimal design of a contract with EAR 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 imposed by the randomization, but is helpful in allowing better diversification of the measurement errors. The next proposition describes how the optimal value of k varies with several parameters of the contracting environment and with the level of aggregate effort to be induced.

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 , given $\delta > \lambda$ (i.e., the stronger the principal's preference for balanced efforts);

would be such that the principal would strictly prefer to choose the schedule rewarding task 2 more (less) highly than task 1.

(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 smaller is $\sigma^2(1 - \rho)$ (i.e., the lower the importance of diversification of the risk from the shocks to measured performance).

Furthermore, the optimal level of k is smaller, the smaller is B (i.e., the smaller the level of aggregate effort to be induced).

In Section 6, where we identify environments where ambiguous schemes outperform deterministic ones, we will build on these results. In Section 7.2, where we study EAR in a setting with n tasks, we show that changes in the number of randomly chosen tasks to reward have the same qualitative effects on incentives and risk as changes in the weighting parameter k in the two-task model, so the comparative statics results for the optimal number of tasks to reward are the same as those above for the optimal k . For now, though, we turn to a second class of ambiguous contracts.

4.2 Ex Post Discretion

A contract involving *ex post discretion* (EPD) also contains two parameters, $\beta > 0$ and $k \in (-1, 1)$. Under such a contract, the principal, *after* observing x_1 and x_2 , chooses whether to pay the agent according to $w = \alpha + \beta x_1 + k\beta x_2$ or $w = \alpha + \beta x_2 + k\beta x_1$. 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 EPD, unlike with EAR, the agent's choice of efforts influences 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 schedules, and if k were equal to 1, EPD would involve no discretion at all and would collapse to the SD contract.

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 agent's optimal effort choices, we use Cain's (1994) derivation of 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. When EPD induces interior solutions for efforts,

(ii) each type of agent chooses effort on his less costly task, \bar{e}^{EPD} , and effort on his more costly task, \underline{e}^{EPD} , satisfying

$$\bar{e}^{EPD} + \lambda \underline{e}^{EPD} = \frac{\beta(1+k)}{\lambda+1} \quad (7)$$

$$\exp[r\beta(1-k)d] \frac{\Phi\left(\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)}{\Phi\left(\frac{-d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)} = \frac{\lambda-k}{1-k\lambda}, \quad (8)$$

where $\theta \equiv \sigma[2(1-\rho)]^{\frac{1}{2}}$ and $d \equiv \bar{e}^{EPD} - \underline{e}^{EPD}$;

(iii) the gap in efforts, $\bar{e}^{EPD} - \underline{e}^{EPD}$, is increasing in λ , approaching 0 as $\lambda \rightarrow 1$; decreasing in $r\beta$, approaching 0 as $r\beta \rightarrow \infty$; increasing in $\sigma^2(1-\rho)$, approaching 0 as $\sigma^2(1-\rho) \rightarrow 0$; and increasing in k , approaching 0 as $k \rightarrow -1^+$;

(iv) the principal's payoff under EPD, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$, is

$$\begin{aligned} \Pi^{EPD}(\beta, k) = & \underline{e}^{EPD} + \frac{1}{\delta} \bar{e}^{EPD} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2} r \sigma^2 \beta^2 (1 + 2\rho k + k^2) \\ & - \frac{1}{r} \ln \{ \exp[-r\beta(1-k)d] \Phi(-) + \Phi(+)\} - \beta(1-k)d\Phi(-d/\theta) + \beta\theta(1-k)\phi(d/\theta), \end{aligned} \quad (9)$$

where $\Phi(-) \equiv \Phi(-d/\theta + r\beta\theta(1-k)/2)$ and $\Phi(+)$ $\equiv \Phi(d/\theta + r\beta\theta(1-k)/2)$.

(v) For any $\lambda > 1$, when EPD and EAR both induce interior solutions for efforts, then $\bar{e}^{EPD} - \underline{e}^{EPD} < \bar{e}^{EAR} - \underline{e}^{EAR}$.

Comparing Propositions 4 and 2 reveals important similarities, as well as important differences, between the incentives and payoffs generated by EPD and EAR. Part (i) holds in each case for exactly the same reason. Moreover, aggregate effort, $\bar{e} + \lambda \underline{e}$, is the same under EPD as under EAR—compare equations (7) and (1).¹⁸ 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, $(\lambda+1)(\bar{e} + \lambda \underline{e})$. Just as for EAR, the first-order conditions for interior optimal efforts then imply equation (4). However, for EPD

$$\frac{E[U'(\cdot)I_{\{x \text{ is rewarded more highly than } \bar{x}\}}]}{E[U'(\cdot)I_{\{\bar{x} \text{ is rewarded more highly than } x\}}]} = \exp[r\beta(1-k)d] \frac{\Phi\left(\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)}{\Phi\left(-\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)},$$

which when combined with equation (4) gives us equation (8).

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, this 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 use means that the more the agent focuses his effort on his preferred task, the less likely that task is to be the more highly rewarded one, so the lower the relative marginal return to that task. Formally, the left-hand side of equation (8), which is increasing in $\bar{e} - \underline{e}$, is strictly greater than the left-hand side of (2) for all $(\bar{e} - \underline{e}) > 0$, and this implies that for all $\lambda > 1$, $\bar{e}^{EPD} - \underline{e}^{EPD} < \bar{e}^{EAR} - \underline{e}^{EAR}$.

Under EPD, the agent's optimal choice of $\bar{e} - \underline{e}$ is smaller the smaller is λ (because choosing balanced efforts is less costly) and the larger is $r\beta$ (because the stronger desire to self-insure is the dominant effect), and as either $\lambda \rightarrow 1$ or $r\beta \rightarrow \infty$, $(\bar{e} - \underline{e}) \rightarrow 0$. These results parallel those for EAR. However, while σ^2 and ρ have no effect on the effort gap 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 $\bar{e} - \underline{e}$ more likely to affect which compensation schedule the principal chooses, so gives the agent a stronger incentive to balance his efforts. As $\sigma^2(1 - \rho) \rightarrow 0$, for example because the shocks become perfectly correlated, optimal efforts become perfectly balanced.

¹⁸Intuitively, we might expect that the principal's freedom, under EPD, to choose the compensation schedule that minimizes his 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, $\bar{e} + \underline{e}$, is lower under EPD than under EAR. Nevertheless, aggregate effort $\bar{e} + \lambda \underline{e}$, and hence the costs of effort incurred, are the same under the two schemes.

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 ($\bar{e} - \underline{e}$) is more complex under EPD than under EAR, nevertheless the effects of k that operate under EAR are the dominant ones under EPD.

What is the cost of the risk imposed by ex post discretion on the agent, and how does it compare to that imposed by ex ante randomization? In the principal's payoff expression (9), the total cost of the risk imposed is given by $\frac{1}{2}r\sigma^2\beta^2(1 + 2\rho k + k^2)$ plus the terms on the second line. We can show that, in fact, EPD with coefficients β and k imposes *lower* risk cost 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 $\frac{1}{2}r\sigma^2\beta^2(1 + 2\rho k + k^2)$.¹⁹ The intuitive reason for this risk-reduction under EPD 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 $\alpha + \beta x_2 + k\beta x_1$. Section 4.1 showed, by contrast, that for any given β and k , EAR imposes *higher* risk cost than would either of these deterministic contracts.

As we did for EAR, we can use $B \equiv \beta(1 + k)$ to derive from (9) an expression for the principal's payoff under EPD as the weighting factor k varies, holding B and hence aggregate effort fixed. As with EAR, increasing k has three distinct effects. By inducing less balanced efforts, it reduces the principal's benefit $\underline{e} + \frac{\bar{e}}{\delta}$. On the other hand, a larger k improves the diversification of the measurement errors. Finally, as k and hence $\bar{e} - \underline{e}$ rises, the extent of the risk reduction from basing pay on $\min\{x_1 + kx_2, x_2 + kx_1\}$ rather than on $x_i + kx_j$ can be shown to decrease. Because the effects of varying k are more complex under EPD than under EAR, it is difficult to generalize all of Proposition 3, but we can show that under EPD, just as under EAR, the optimal level of k , for any given level of aggregate effort, will be smaller, the larger is the principal's complementarity parameter δ .

4.3 Ex Ante Randomization versus Ex Post Discretion

The preceding paragraphs have argued that, for any (β, k) that induce interior solutions for efforts under both EAR and EPD, (i) EPD induces a strictly smaller gap in efforts $\bar{e} - \underline{e}$ than EAR, while the two schemes induce the same aggregate effort $\bar{e} + \lambda\underline{e}$ and hence the same total cost of effort, and (ii) EPD imposes lower risk costs on the agent than EAR. Since whenever $\delta \geq \lambda$, balanced efforts are socially efficient, taken together the findings above yield:

Proposition 5 *If, for given β and $k \in (-1, \frac{1}{\lambda})$, EAR and EPD induce interior solutions for efforts, and if $\delta \geq \lambda$, then EPD generates at least as great a payoff for the principal as EAR.*

5 When Are Deterministic Contracts Optimal?

This section identifies three environments in which both types of ambiguous incentive scheme are strictly dominated by a deterministic contract.

¹⁹This is proved formally in Step 1 of the proof of Proposition 5.

Proposition 6 For given (β, k) , where $k \in (-1, 1)$, both EAR and EPD yield strictly lower payoff for the principal than a suitably designed symmetric deterministic (SD) contract, if any of the following conditions holds:

- (i) $\lambda = 1$ and $\rho < 1$;
- (ii) EAR and EPD induce the agent to exert effort only on his preferred task;
- (iii) $\delta < \lambda$.

Underlying this proposition is the result that, for any $k < 1$, and for any λ , a SD contract can induce any given level of *aggregate effort* $\bar{e} + \lambda e$ with the imposition of lower risk costs on the agent than EAR or EPD. At the same time, whenever $\lambda > 1$, a SD contract 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. If $\lambda = 1$, the SD contract and the two ambiguous schemes all induce perfectly balanced efforts. The proposition shows that, in this case, the SD contract imposes strictly lower risk than does either EAR or EPD, as long as $\rho < 1$. With EAR this result is clear, since even though the balanced efforts eliminate the risk from the randomization, the fact that $k < 1$ means that EAR weights the performance measures unequally, while the SD contract weights them equally and so better diversifies the risk from the shocks. That EPD, while inducing the same efforts as the SD contract, imposes strictly more risk is less obvious, but we show that the cost of risk imposed by EPD, when $\lambda = 1$, is strictly decreasing in k , when β is adjusted to keep aggregate effort fixed.²⁰ Since the SD contract corresponds to $k = 1$, the SD contract therefore strictly dominates EPD.

Under condition (ii), even the ambiguous schemes induce corner solutions. This happens, for example, when, holding $r\sigma^2$ fixed, r gets sufficiently small (so the self-insurance motive becomes sufficiently weak) and σ^2 gets sufficiently large (so under EPD, a change in the agent's efforts is sufficiently unlikely to affect the ex post choice of compensation schedule). Under condition (iii), the socially efficient effort allocation involves fully focused efforts; hence, for any fixed level of aggregate effort and thus any fixed cost of effort incurred, a shift towards more balanced efforts would actually reduce the principal's payoff. Therefore, under any of conditions (i), (ii), or (iii), 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.

6 When Are Ambiguous Contracts Optimal?

We now identify three environments in which ambiguous contracts, when designed optimally, dominate the best deterministic scheme. In each of these environments, both ex ante randomization and ex post discretion, with the parameter k adjusted optimally, induce the agent to choose *perfectly* balanced efforts, and EAR is as attractive for the principal as EPD. In all three environments, we show that optimally

²⁰The proof shows that as k is increased, the benefit of reducing the variance of $(x_i + kx_j)/(1+k)$ outweighs the cost of increasing the correlation between $(x_1 + kx_2)/(1+k)$ and $(x_2 + kx_1)/(1+k)$.

weighted EAR and EPD generate a payoff for the principal arbitrarily close to that he could achieve if he knew the agent's preferences across tasks, so ambiguous schemes eliminate the efficiency losses from the agent's hidden information.

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

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 λ is strictly greater than but arbitrarily close to 1, which we term the limiting case as $\lambda \rightarrow 1^+$.

For both EAR and EPD, we saw in Section 4 that the agent's effort choices and the principal's payoff are continuous in λ at $\lambda = 1$. In contrast, Section 3 showed that the principal's payoff from the best deterministic scheme drops discontinuously as λ is increased from 1, whenever δ is such that the best deterministic scheme is an ADM. The robustness of EAR and EPD to the introduction of hidden information underlies the superiority of the ambiguous schemes in the limiting case as $\lambda \rightarrow 1^+$, as we now show.

Propositions 2 and 4 show that as $\lambda \rightarrow 1$, so the two tasks become equally costly, $(\bar{e} - \underline{e}) \rightarrow 0$ for any $k \in (-1, 1)$, under both EAR and EPD. Equations (5) and (6) show how varying k affects the principal's payoff from EAR, $\Pi^{EAR}(B, k)$, holding fixed at $\frac{B}{1+\lambda}$ the level of aggregate effort induced. Whereas in general, as discussed in Section 4.1, increasing k has conflicting effects on $\Pi^{EAR}(B, k)$, in the limit as $\lambda \rightarrow 1$, the situation is dramatically simpler:

$$\lim_{\lambda \rightarrow 1} \Pi^{EAR}(B, k) = \frac{(\delta + 1)B}{\delta} \frac{1}{4} - \frac{B^2}{8} - \frac{1}{2} r \sigma^2 B^2 \left(\frac{1 + 2\rho k + k^2}{(1 + k)^2} \right). \quad (10)$$

As $\lambda \rightarrow 1$, efforts under EAR become perfectly balanced—both \bar{e} and \underline{e} approach $\frac{B}{4}$ —and therefore the risk cost imposed by the random choice of compensation schedule tends to zero. Hence an increase in k has only one effect on $\Pi^{EAR}(B, k)$, holding B fixed: it improves the diversification of the shocks to measured performance, as reflected in the final term of (10). Thus, as $\lambda \rightarrow 1$, $\Pi^{EAR}(B, k)$ is increasing in k as long as k induces interior solutions, which it does as long as $k < \frac{1}{\lambda}$. Therefore, as $\lambda \rightarrow 1$, $\Pi^{EAR}(B, k)$ is maximized, for any B , by setting k arbitrarily close to, but less than, 1 ($k \rightarrow 1^-$). With k set in this way, the principal's payoff approaches

$$\lim_{k \rightarrow 1} \lim_{\lambda \rightarrow 1} \Pi^{EAR}(B, k) = \frac{(\delta + 1)B}{\delta} \frac{1}{4} - \frac{B^2}{8} - \frac{1}{4} r \sigma^2 B^2 (1 + \rho). \quad (11)$$

The right-hand side of (11) equals the payoff the principal would achieve, if λ were exactly equal to 1, from a symmetric deterministic (SD) contract with $\beta = \frac{B}{2}$, since such a contract would induce effort $\frac{B}{4}$ on each task and generate the same diversification of the shocks as EAR does when $k \rightarrow 1^-$.²¹ Thus, as $\lambda \rightarrow 1^+$, the principal's payoff under optimally weighted EAR gets arbitrarily close to that from a SD contract when the agent has no preference between tasks.

Proposition 5 implies that for any $k < 1$ and for any B , in the limit as $\lambda \rightarrow 1^+$, EPD must yield the principal a payoff at least as large as $\Pi^{EAR}(B, k)$. Part (i) of Proposition 6 showed that the principal's payoff under EPD, for any $k < 1$ and for any B , is, even in the limit as $\lambda \rightarrow 1^+$, less than his payoff under the SD contract inducing the same levels of effort. As noted above, such a SD contract has $\beta = \frac{B}{2}$

²¹See equation (23) in the proof of Lemma 1 in the appendix.

and generates a payoff given by the right-hand side of (11). It therefore follows that as $\lambda \rightarrow 1^+$, optimally designed EPD, like EAR, sets k arbitrarily close to, but less than, 1 ($k \rightarrow 1^-$), and yields a payoff arbitrarily close to that from a SD contract at $\lambda = 1$.

Section 3.2 shows that in the no hidden information benchmark, the efforts and payoff from the contract pair (C_1^{bal}, C_2^{bal}) are continuous at $\lambda = 1$, where they match the efforts and payoff from the SD contract. On the other hand, Section 3.3 shows that under hidden information, even as λ gets arbitrarily close to 1, it is impossible for the principal to induce balanced efforts from both types of agent. Consequently, the payoff from an asymmetric deterministic menu (ADM), even as $\lambda \rightarrow 1^+$, remains strictly below that from a SD contract at $\lambda = 1$, whenever $\delta > \lim_{\lambda \rightarrow 1} \delta^{HI}(\lambda, r\sigma^2, \rho) = \lim_{\lambda \rightarrow 1} \delta^{NHI}(\lambda, r\sigma^2, \rho)$. The arguments in the preceding paragraphs together imply:

Proposition 7 *Consider the limiting case as $\lambda \rightarrow 1^+$. Under both EAR and EPD, for any given level of aggregate effort, $\bar{e} + \lambda \underline{e}$, to be induced:*

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

(ii) *the optimal value of $k \rightarrow 1^-$;*

(iii) *with k adjusted optimally, the principal's payoff under both ambiguous schemes approaches his payoff in the no hidden information benchmark from (C_1^{bal}, C_2^{bal}) . This limiting payoff equals the principal's payoff from the symmetric deterministic (SD) contract at $\lambda = 1$.*

Therefore, for $\delta > \lim_{\lambda \rightarrow 1} \delta^{NHI}(\lambda, r\sigma^2, \rho)$, EAR and EPD with k adjusted optimally dominate the best deterministic scheme under hidden information.

The preceding analysis also has another implication. Even if $\lambda = 1$, the symmetric deterministic contract leaves the agent indifferent over how any total effort is split between the tasks, while under EAR and EPD, for any $k < 1$, the optimal allocation of efforts is unique. Thus when $\lambda = 1$, with the weighting parameter k set arbitrarily close to, but less than, 1, EAR and EPD not only generate a payoff for the principal arbitrarily close to the theoretical payoff from the SD contract, but they also ensure that the agent has a strict preference for choosing perfectly balanced efforts.

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

Consider now the case where, holding $r\sigma^2$ fixed at $R < \infty$, r gets very large and σ^2 very small. Propositions 2 and 4 show that, in this environment, for any $k \in (-1, \frac{1}{\lambda})$, $(\bar{e} - \underline{e}) \rightarrow 0$ under both EAR and EPD: As the agent becomes infinitely risk-averse, it becomes optimal for him under both types of ambiguous scheme to choose perfectly balanced efforts, so providing full self-insurance against the uncertainty over which compensation schedule will ultimately be used. In contrast, as long as $r\sigma^2$ remains unchanged, the efforts induced and the payoff generated by all deterministic schemes are unaffected.

Under EAR, in the limit as $r \rightarrow \infty$ and $\sigma^2 = \frac{R}{r} \rightarrow 0$, both \bar{e} and \underline{e} approach $\frac{B}{(\lambda+1)^2}$ (as long as $k < \frac{1}{\lambda}$), and with perfectly balanced efforts, the risk cost imposed on the agent by the randomization tends to zero. As a consequence, $\Pi^{EAR}(B, k)$, as given by equations (5) and (6), simplifies to

$$\lim_{r \rightarrow \infty, \sigma^2 = R/r \rightarrow 0} \Pi^{EAR}(B, k) = \frac{(\delta + 1)}{\delta} \frac{B}{(\lambda + 1)^2} - \frac{B^2}{2(\lambda + 1)^2} - \frac{1}{2} RB^2 \frac{1 + 2\rho k + k^2}{(1 + k)^2}. \quad (12)$$

Under EPD, the principal's payoff for given B and k approaches the same value as under EAR as $r \rightarrow \infty$ and $\sigma^2 = \frac{R}{r} \rightarrow 0$:²²

$$\lim_{r \rightarrow \infty, \sigma^2 = R/r \rightarrow 0} \Pi^{EPD}(B, k) = \lim_{r \rightarrow \infty, \sigma^2 = R/r \rightarrow 0} \Pi^{EAR}(B, k). \quad (13)$$

Equations (12) and (13) show that, when r gets very large and $\sigma^2 = \frac{R}{r}$ very small, the only effect on $\Pi^{EAR}(B, k)$ and $\Pi^{EPD}(B, k)$ of increasing k , over the range $k \in (-1, \frac{1}{\lambda})$ where the induced gap in efforts ($\bar{e} - \underline{e}$) is approximately 0, is to improve the diversification of the shocks to measured performance. Hence, under both EAR and EPD, it is optimal to set k arbitrarily close to, but less than, $\frac{1}{\lambda}$ ($k \rightarrow (\frac{1}{\lambda})^-$), and doing so generates for the principal a payoff approaching

$$\lim_{k \rightarrow (1/\lambda)^-} \lim_{r \rightarrow \infty, \sigma^2 = R/r \rightarrow 0} \Pi^{EAR}(B, k) = \frac{(\delta + 1)}{\delta} \frac{B}{(\lambda + 1)^2} - \frac{B^2}{2(\lambda + 1)^2} - \frac{1}{2} RB^2 \frac{\lambda^2 + 2\rho\lambda + 1}{(\lambda + 1)^2}. \quad (14)$$

The right-hand side of (14) is exactly the payoff the principal would obtain, in the NHI benchmark, from using (C_1^{bal}, C_2^{bal}) with $\beta = \frac{B}{1+\lambda}$, since this pair of contracts would induce from each type of agent effort $\frac{B}{(\lambda+1)^2}$ on each task and would impose a risk premium given by the final term.²³

Since as $r \rightarrow \infty$ and $\sigma^2 = \frac{R}{r} \rightarrow 0$, optimally weighted EAR and EPD allow the principal to get arbitrarily close to the payoff he would receive in the NHI benchmark, and since Proposition 1 shows that under hidden information, whenever $\delta > \delta^{NHI}$, the best deterministic scheme leaves the principal strictly worse off than he would be in the NHI benchmark, we have proved:

Proposition 8 *Consider the limiting case where $r \rightarrow \infty$ and $\sigma^2 = \frac{R}{r} \rightarrow 0$. Under both EAR and EPD, for any given level of aggregate effort, $\bar{e} + \lambda \underline{e}$, to be induced:*

- (i) *the gap in efforts, $\bar{e} - \underline{e}$, approaches 0 for any λ and for any $k < \frac{1}{\lambda}$;*
- (ii) *the optimal value of $k \rightarrow (\frac{1}{\lambda})^-$;*
- (iii) *with k adjusted optimally, the principal's payoff under both ambiguous schemes approaches his payoff in the no hidden information benchmark from (C_1^{bal}, C_2^{bal}) .*

Therefore, for $\delta > \delta^{NHI}(\lambda, R, \rho)$, EAR and EPD with k adjusted optimally dominate the best deterministic scheme under hidden information.

6.3 The Limiting Case of Perfect Correlation of the Shocks

Under ex post discretion, as the correlation between the shocks to measured performance on the two tasks approaches 1, the optimal efforts chosen by the agent become perfectly balanced ($\bar{e} - \underline{e} \rightarrow 0$) for any $k < \frac{1}{\lambda}$, as shown by Proposition 4. This reflects the fact that for any effort choices (\bar{e}, \underline{e}) , as $\sigma^2(1 - \rho) \rightarrow 0$, the agent's uncertainty about which compensation schedule the principal will choose

²²This can be proved by first substituting $B = \beta(1 + k)$ into equation (9) and then noting that the terms on the second line all tend to 0 as $r \rightarrow \infty$ and $\sigma^2 = \frac{R}{r} \rightarrow 0$. The equality in (13) reflects the fact that not only do the efforts under EPD and EAR approach the same values, but the cost of the risk imposed by the two schemes becomes equal, because as $\sigma^2 \rightarrow 0$, $\min\{x_1 + kx_2, x_2 + kx_1\}$ becomes no less variable than either $x_1 + kx_2$ or $x_2 + kx_1$ alone, and thus the total cost of risk under EPD is given by the final term in (12).

²³See equation (26) in the proof of Lemma 2 in the appendix.

disappears. For any (B, k) with $k < \frac{1}{\lambda}$, \bar{e} and \underline{e} approach $\frac{B}{(\lambda+1)^2}$, and $\Pi^{EPD}(B, k)$ approaches

$$\lim_{\rho \rightarrow 1} \Pi^{EPD}(B, k) = \frac{\delta + 1}{\delta} \frac{B}{(\lambda + 1)^2} - \frac{B^2}{2(\lambda + 1)^2} - \frac{1}{2} r \sigma^2 B^2. \quad (15)$$

This limiting payoff is independent of k as long as $k < \frac{1}{\lambda}$: As the shocks to measured performance become perfectly correlated, any possibility of diversifying the risk from the shocks, by increasing k , disappears, and hence the risk cost of EPD reduces to $\frac{r\sigma^2 B^2}{2}$. Since (15) is independent of k , any value of k in $(-1, \frac{1}{\lambda})$ is optimal under EPD as $\rho \rightarrow 1$. Furthermore, (15) matches what the principal would obtain, in the NHI benchmark with $\rho = 1$, from using (C_1^{bal}, C_2^{bal}) to induce perfectly balanced efforts and setting $\beta = \frac{B}{1+\lambda}$.²⁴ Thus, in this limiting environment as well, EPD allows the principal to achieve a payoff as high as he would achieve in the absence of hidden information.

Under EAR, too, as $\rho \rightarrow 1$, diversification of the risk from the shocks becomes impossible, so in equation (5) the cost of risk due to the shocks approaches $\frac{r\sigma^2 B^2}{2}$. However, as Proposition 2 shows, under EAR, in contrast to EPD, the incentive for balanced efforts is independent of $\sigma^2(1 - \rho)$, since the realized outputs have no effect on the random choice of compensation schedule. Nevertheless, Proposition 2 also showed that lowering the weighting parameter k lowers the induced gap in efforts $\bar{e} - \underline{e}$. In the limit as $\rho \rightarrow 1$, equations (5) and (6) show that $\Pi^{EAR}(B, k)$ is decreasing in k , for any fixed B : Lowering k does not affect the risk premium due to the shocks but, by reducing the effort gap, does raise the principal's benefit as well as reduce the risk cost of the exogenous randomization. Hence as $\rho \rightarrow 1$, it is optimal under EAR, for any level of aggregate effort to be induced, to set k arbitrarily close to, but larger than, -1 ($k \rightarrow -1^+$), thus inducing a gap in efforts arbitrarily close to, but larger than, 0 (as shown by (2)). With k set in this way, the principal achieves under EAR a payoff arbitrarily close to the right-hand side of (15).²⁵

Thus, as the correlation of the shocks approaches 1, optimally weighted EAR and EPD both allow the principal to approach the payoff he would receive in the NHI benchmark. Since Proposition 1 shows that under hidden information, whenever $\delta > \delta^{NHI}$, the best deterministic scheme leaves the principal strictly worse off than he would be in the NHI benchmark, we have shown:

Proposition 9 *Consider the limiting case of perfect correlation of the shocks: $\rho \rightarrow 1$. For any given level of aggregate effort, $\bar{e} + \lambda \underline{e}$, to be induced:*

(i) *under EPD, the gap in efforts, $\bar{e} - \underline{e}$, 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, $\bar{e} - \underline{e}$, approaches 0 for any λ as $k \rightarrow -1^+$, and the optimal value of $k \rightarrow -1^+$;*

(iii) *with k adjusted optimally, the principal's payoff under both ambiguous schemes approaches his payoff in the no hidden information benchmark from (C_1^{bal}, C_2^{bal}) .*

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.

Analogous arguments and conclusions hold in the limiting environment where the variance σ^2 of the

²⁴See equation (26) in the appendix.

²⁵As k is lowered, the coefficient β must be raised to keep aggregate effort, which is proportional to $B \equiv \beta(1 + k)$, fixed. The value of k must remain slightly larger than -1 to ensure that aggregate effort is strictly positive.

shocks to measured performance approaches 0. In this environment, optimally weighted EAR and EPD allow the principal to get arbitrarily close to the outcome in which perfectly balanced efforts are induced with the imposition of no risk cost on the agent, so the ambiguous schemes allow the principal to achieve a payoff arbitrarily close to what he would achieve in the absence of *any* informational asymmetries. The best deterministic scheme under hidden information cannot achieve balanced efforts, even as $\sigma^2 \rightarrow 0$, so the best deterministic scheme is dominated by optimally weighted ambiguous ones, for $\delta > \delta^{NHI}(\lambda, 0, \rho) = \lambda$.

6.4 Discussion

We have identified three environments in which ambiguous contracts, when designed optimally, dominate the best deterministic scheme. In each of these environments, both EAR and EPD, with the parameter k adjusted optimally, 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 his payoff in the no hidden information benchmark from inducing perfectly balanced efforts, so in these environments, ambiguous incentive schemes eliminate the efficiency losses from the agent's hidden information. In each of these settings, optimally designed ambiguous schemes dominate the best deterministic scheme whenever, in the NHI benchmark, inducing balanced efforts is optimal, that is, whenever $\delta > \delta^{NHI}(\lambda, r\sigma^2, \rho)$. If, instead, $\delta < \delta^{NHI}(\lambda, r\sigma^2, \rho)$, then the principal's benefit from balanced efforts is outweighed by the extra risk costs imposed by ambiguous schemes, and a SDM, inducing fully focused efforts but providing better insurance, is optimal.

In none of these settings is the asymmetric deterministic menu (ADM) ever optimal for the principal. There are two reasons for this result. First, because the ADM induces one type of agent to choose perfectly balanced efforts and the other type fully focused ones, its outcome is roughly intermediate between the ambiguous schemes and the SDM. Second, whenever $\lambda > 1$, the ADM leaves informational rents to one type of agent, whereas neither the ambiguous schemes nor the SDM leave any rents to either type. These rents make the ADM strictly less profitable, for all values of δ , than the best alternative scheme.

7 Extensions and Robustness

7.1 Linear Ambiguous Contracts vs. Nonlinear Contracts

Our analysis has shown that ex ante randomization and ex post discretion succeed in mitigating the privately-informed agent's incentives for gaming even under the restriction that ex post, the agent's compensation schedule is linear and separable in the performance measures. In contrast, linear, separable deterministic schemes are unable to induce both types of privately-informed agents to choose strictly positive efforts on both tasks. These contrasting results raise the question of the relationship between our two classes of ambiguous contracts and nonlinear, nonseparable schemes.

In our view, the mechanism by which ex ante randomization alleviates incentives for gaming is one that is familiar and readily comprehensible. Because the two possible compensation schedules that may ultimately be used are linear, they are simple to describe and implement. Furthermore, the desirability of spreading one's efforts across tasks as a means of reducing the riskiness of compensation under ex

ante randomization is analogous to strategies for diversifying risk in many familiar settings. For example, Bevan and Hood (2004), in the context of performance measurement in healthcare, argue for making the relative weights on the measures unpredictable by invoking the “analogy [...] with the use of unseen examinations, where the unpredictability of what the questions will be means that it is safest for students to cover the syllabus” (p.598).

Consider, by way of contrast, the following nonlinear deterministic compensation scheme:

$$w(x_1, x_2) = \alpha - \frac{1}{r} \ln \left\{ \frac{1}{2} \exp[-r\beta(x_1 + kx_2)] + \frac{1}{2} \exp[-r\beta(x_2 + kx_1)] \right\} \quad (16)$$

Faced with such a scheme, how would an agent figure out what allocation of effort across tasks was most beneficial? He might be able to work out that the marginal impact on pay of an increase in x_1 was decreasing in the level of x_1 and increasing in the level of x_2 (with symmetric results for increases in x_2), and hence he might be able to deduce that such a scheme, relative to a symmetric linear and separable one, discourages focused and encourages balanced effort allocations. In fact, for any (α, β, k) , this nonlinear, nonseparable deterministic scheme provides the same expected utility for an agent, as a function of the effort choices, as does ex ante randomization. Moreover, because the wage is a deterministic function of (x_1, x_2) , this scheme in theory would generate a higher expected payoff for the principal than does EAR. But such a conclusion rests on the assumption that an agent would choose the same efforts when faced with (16) as when faced with EAR. We contend that EAR provides considerably more *salient* incentives for balanced efforts than does (16) and that, as a result, the incentive effects of ex ante randomization are likely to be more consistent and more predictable than those of the complicated deterministic scheme given in (16).²⁶

It is also worth noting that in the three environments studied in Section 6, the incentive scheme in (16), even if fully understood by the agent, would perform no better than ex ante randomization, since in each of these settings, optimally weighted EAR induces full self-insurance and so generates no extra risk costs from randomization. More generally, since *ceteris paribus*, EAR is attractive in environments where it generates very strong incentives for balanced efforts, when EAR is attractive it will generate a payoff for the principal close to the theoretical payoff from the scheme in (16).

Consider now ex post discretion, under which the choice between two linear compensation schedules is made strategically by the principal, rather than randomly. Again, the ex post linearity of the schedules makes them simple to describe and implement, and we contend that the incentives EPD provides for balanced efforts are also very salient, since an agent will recognize that the principal will choose to pay him based on whichever of $x_1 + kx_2$ and $x_2 + kx_1$ is smaller. Perhaps equally important, while in our model EPD achieves the same outcome as if the principal could commit to the nonlinear, nonseparable wage scheme $w = \min\{\alpha + \beta x_1 + k\beta x_2, \alpha + \beta x_2 + k\beta x_1\}$, EPD requires less commitment power on the principal’s part. This has two implications. First, in conjunction with Propositions 4 and 5, it implies that the beneficial incentive effects of EAR are robust even if the agent suspects that the principal might deviate to EPD, by waiting to observe outputs before choosing between the two compensation schedules. Second, in a more general version of our model, allowing the principal ex post discretion over the choice of compensation schedule might allow him to do strictly better than he would if forced to commit ex ante to a wage scheme. In such a model,

²⁶Englmaier, Roider, and Sunde (2012) and Chetty, Looney, and Kroft (2009) provide evidence from field experiments of the significant impact of changes in salience of incentive schemes and of taxes, respectively.

the principal would be ex ante uncertain about the efficient effort profile and ex post better informed, so ex post discretion would not only provide the incentive benefits our analysis highlighted, but would also allow the principal to reward the type of performance which in the circumstances turned out to be appropriate.²⁷

7.2 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. 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 the dimensions of performance. We now study some of the trade-offs 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 $(\epsilon_1, \dots, \epsilon_n)$ 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 $c_i(e_1, \dots, e_n) = \frac{1}{2}(\lambda e_i + \sum_{j \neq i} e_j)^2$. 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 payoff be given by

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

where δ parameterizes the strength of the principal's desire for a balanced effort profile. As in the two-task model, the socially efficient effort profile is perfectly balanced whenever $\delta > \lambda$.

Consider 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 trade-off between the effects on incentives and risk 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 \bar{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 $(\lambda \underline{e} + (n-1)\bar{e})$ and the gap in efforts $\bar{e} - \underline{e}$ satisfy, respectively,

$$\lambda \underline{e} + (n-1)\bar{e} = \frac{\kappa\beta}{n-1+\lambda} \quad \text{and} \quad \bar{e} - \underline{e} = \frac{1}{r\beta} \ln \left[\frac{\lambda(n-\kappa)}{(n-1) - (\kappa-1)\lambda} \right]. \quad (17)$$

Reducing κ , the number of tasks rewarded, increases the agent's uncertainty about the level of compensation he will receive and so strengthens his incentive to self-insure. As a result, the agent's optimal effort profile is more balanced ($\bar{e} - \underline{e}$ 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 (17), we can write the principal's

²⁷Scott and Triantis (2006) and Prendergast (1999) develop related arguments.

payoff as a function of $\tilde{\beta}$ and κ , when α is set to ensure zero rents for each type of agent:

$$\begin{aligned} \Pi(\tilde{\beta}, \kappa) = \underline{e} + \frac{(n-1)}{\delta} \bar{e} - \frac{\tilde{\beta}^2}{2(n-1+\lambda)^2} \\ - \frac{1}{2} r \sigma^2 \tilde{\beta}^2 \frac{1+\rho(\kappa-1)}{\kappa} - \frac{1}{nr} \ln \left[\frac{(n-\kappa)^{n-\kappa} (n-1+\lambda)^n}{n^n \lambda^\kappa ((n-1) - (\kappa-1)\lambda)^{n-\kappa}} \right], \end{aligned} \quad (18)$$

where

$$\underline{e} + \frac{n-1}{\delta} \bar{e} = \frac{\delta+n-1}{\delta} \frac{\tilde{\beta}}{(n-1+\lambda)^2} - \frac{(\delta-\lambda)(n-1)\kappa}{\delta(n-1+\lambda)r\tilde{\beta}} \ln \left[\frac{\lambda(n-\kappa)}{(n-1) - (\kappa-1)\lambda} \right]. \quad (19)$$

Holding $\tilde{\beta}$ fixed and varying κ isolates the effect of changing the number of tasks rewarded, holding fixed the level of aggregate effort. Comparison of equations (18)-(19) with equations (5)-(6) 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 , by inducing a larger gap $\bar{e} - \underline{e}$, has two negative effects: i) it lowers the principal's benefit $\underline{e} + \frac{n-1}{\delta} \bar{e}$ when aggregate effort is held fixed, as long as $\delta > \lambda$ (this corresponds to the fact that (19) is decreasing in κ), and ii) it raises the cost of compensating the agent for the risk imposed by the randomization (this corresponds to the fact that the term in square brackets in (18) is increasing in κ). At the same time, raising κ also improves the diversification of the risk from the shocks to measured performance, thereby reducing the risk premium per unit of aggregate effort induced (as reflected in the fact that $\frac{1+\rho(\kappa-1)}{\kappa}$ in (18) is decreasing in κ).

Given the qualitative similarity between the role of κ in the n -task model and that of k in the two-task model, it is relatively straightforward (see the appendix) to derive the following comparative statics results for the optimal number of tasks to reward, given any desired level of aggregate effort to be induced. Analogously with Proposition 3, the optimal number of tasks to reward is smaller, i) the stronger is the principal's preference for balanced efforts; ii) the less risk-averse the agent is, holding $r\sigma^2$ fixed; and iii) the lower the importance of diversification of the risk from the shocks to measured performance (i.e. the lower is $\sigma^2(1-\rho)$). In addition, the weaker the incentives for aggregate effort, the smaller the optimal number of tasks to reward in inducing that effort.

7.3 Menus of Incentive Schemes with Ex Ante Randomization

This section examines whether the performance of ex ante randomization can be improved by the use of menus. Consider the following, incentive-compatible, menu of two incentive schemes each involving randomization. For $k \in (-1, 1)$, Scheme $i \in \{1, 2\}$, intended for the agent who prefers task i , specifies that with probability $p \in (\frac{1}{2}, 1)$, $w = \alpha + \beta x_i + k\beta x_j$, and with probability $1-p$, $w = \alpha + \beta x_j + k\beta x_i$. As $p \rightarrow 1/2$, the two schemes become identical, so the menu reduces to EAR.

The value of p has no effect on aggregate effort. However, as p rises, each type of agent faces less uncertainty about his compensation schedule, hence has weaker incentives to self-insure by balancing his effort choices, so the induced effort gap $\bar{e} - \underline{e}$ rises. In this respect, a larger p mirrors the effect of a larger weighting parameter k . Nevertheless, there is a crucial difference between p and k . An increase in k improves the diversification of the risk from the shocks to measured performance. However, because, regardless of the value of p , the agent is ultimately paid either $\alpha + \beta x_1 + k\beta x_2$ or $\alpha + \beta x_2 + k\beta x_1$,

changes in p have no effect on the diversification of this risk.

In consequence, whereas Proposition 3 and Section 6 showed that the weighting factor k is a valuable instrument in the design of ambiguous schemes, we have the following negative conclusion for the role of p : If a symmetric menu of randomized schemes with parameters (β, k, p) induces interior solutions for efforts, then as long as $\delta > \lambda$, the principal's payoff will be increased by lowering p to $1/2$, thus replacing the menu with EAR as analyzed in Section 4.1. Hence combining menus with ex ante randomization brings no additional benefit for the principal.

7.4 Imperfect Substitutability of Efforts for the Agent

So far we have focused on the case where efforts are perfect substitutes in the agent's cost function. Although this assumption does not qualitatively affect the performance of EAR and EPD, it simplifies the characterization of the optimal linear deterministic scheme. Our key findings are robust to imperfect substitutability of efforts. Specifically, it remains true that i) if tasks are sufficiently complementary for the principal, ambiguous schemes are superior to deterministic ones in settings where they generate very strong incentives for balanced efforts, and ii) in such settings, ambiguous schemes eliminate the efficiency losses from the agent's hidden information.

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

$$c(\bar{e}, \underline{e}) = \frac{1}{2} (\bar{e}^2 + 2s\lambda\bar{e}\underline{e} + \lambda^2\underline{e}^2), \quad (20)$$

where the parameter $s \in [0, 1]$ measures the degree of substitutability of efforts. Perfect substitutability corresponds to $s = 1$ and no substitutability to $s = 0$. To simplify the analysis of deterministic schemes, we will focus on the case where $s \geq \frac{1}{\lambda}$, representing a situation of high, but imperfect, substitutability, and we will let $\delta \rightarrow \infty$.

In the “no hidden information benchmark”, the principal will offer each type of agent a contract of the form $w = \alpha + \beta\bar{x} + v\beta\underline{x}$ with $v \geq 1$, where \bar{x} (respectively, \underline{x}) denotes measured performance on the preferred (respectively, other) task. The weighting parameter v 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 to choose equal efforts on the two tasks: $v^{NHI} = \frac{\lambda(\lambda+s)}{1+s\lambda}$.

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 efforts on both tasks. The optimal deterministic menu is an ADM inducing one type to choose (relatively) balanced efforts while the other type chooses fully focused efforts. We can derive an upper bound on the principal's payoff from an ADM and show that this bound is strictly less than his payoff in the NHI benchmark.

Importantly, the incentives provided by ex ante randomization and ex post discretion are not qualitatively affected by whether efforts are imperfect or perfect substitutes for the agent. EAR continues to give the risk-averse agent an incentive to partially self-insure by choosing relatively balanced efforts on the two tasks, and EPD continues to give even stronger incentives for balance because of the agent's ability to influ-

ence, through his efforts, which task is more highly rewarded. Interior optimal efforts under EAR satisfy

$$\frac{\partial c}{\partial \bar{e}} + \frac{\partial c}{\partial \underline{e}} = \beta(1 + k) \quad (21)$$

and

$$\exp[r\beta(1 - k)(\bar{e} - \underline{e})] = \frac{\frac{c_2}{c_1} - k}{1 - k\frac{c_2}{c_1}}, \quad (22)$$

where $\frac{c_2}{c_1} \equiv \frac{\partial c/\partial \underline{e}}{\partial c/\partial \bar{e}} = \frac{s\lambda\bar{e} + \lambda^2\underline{e}}{\bar{e} + s\lambda\underline{e}}$. Equation (22) is a generalized version of (2) in which the constant λ is replaced by the function $\frac{\partial c/\partial \underline{e}}{\partial c/\partial \bar{e}}$. Optimal efforts under EPD satisfy (21) and equation (8), except that the right-hand side of (8) is replaced by the right-hand side of (22).

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

In these limiting environments, optimally weighted 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 scheme can induce both types of agent to work on both tasks, and as a consequence, the payoff from the best deterministic scheme is bounded away from that 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.

7.5 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 more robust to hidden information of the agent apply even outside the exponential-normal framework. Let the measurement technology remain $x_i = e_i + \varepsilon_i$, but now let $(\varepsilon_1, \varepsilon_2)$ have an arbitrary symmetric joint density. Let each type of agent's utility be $U(w - c(\bar{e}, \underline{e}))$, with $c(\bar{e}, \underline{e})$ as in (20) and $U(\cdot)$ an arbitrary strictly concave function.

Under both EAR and EPD, interior optimal effort choices for each type of agent satisfy

$$\frac{\partial c}{\partial \bar{e}} + \frac{\partial c}{\partial \underline{e}} = \beta(1 + k) \quad \text{and} \quad \frac{E[U'(\cdot)I_{\{\underline{x} \text{ is more highly rewarded}\}}]}{E[U'(\cdot)I_{\{\bar{x} \text{ is more highly rewarded}\}}]} = \frac{\frac{c_2}{c_1} - k}{1 - k\frac{c_2}{c_1}}.$$

Just as for the exponential-normal model, EAR gives the risk-averse agent an incentive to choose more balanced efforts to partially self-insure, and EPD introduces a second motive for balance because the agent's efforts actually influence which compensation schedule is chosen by the principal. However, here

the strength of the insurance motive is not generally equal under the two types of ambiguous schemes, so EPD does not necessarily induce more balanced efforts than EAR.

Nevertheless, whenever the symmetric deterministic contract induces interior efforts, both EAR and EPD do so as well, and effort choices under both EAR and EPD are more balanced than under the SD contract. Furthermore, when efforts are perfect substitutes for the agent ($s = 1$), as λ increases from 1, $\bar{e}^{EAR}/\underline{e}^{EAR}$ and $\bar{e}^{EPD}/\underline{e}^{EPD}$ both increase continuously from 1, whereas $\bar{e}^{SD}/\underline{e}^{SD}$ jumps from 1 to ∞ . Thus, even outside the exponential-normal framework, the performance of both types of ambiguous scheme remains more robust to hidden information.

8 Conclusion

Gaming of incentive schemes is of serious concern to incentive designers in a wide range of settings in both the private and public sectors. We have discussed examples from the UK National Health Service, the US healthcare sector, internet search, and US law school rankings. In all of these examples, the incentive designer cares about the performance of agents along several different dimensions, and deterrence of gaming is hampered by the agents' superior knowledge of the environment. These examples are also ones where a lack of transparency—deliberate ambiguity about the criteria upon which rewards will be based—has been used or advocated to deter gaming.

To elucidate the benefits and costs of ambiguous incentive schemes, we have analyzed a model in which an agent with superior knowledge of the contracting environment—here, his cost of effort on different tasks—can game an incentive scheme by focusing his effort on the task he finds less costly. We have studied two simple types of ambiguous contracts, ex ante randomization (EAR) and ex post discretion (EPD), and have shown how and to what extent they can mitigate this gaming. When the agent has private information about his preferences, the principal in general faces a trade-off between the stronger incentives for balanced efforts under ambiguous schemes and the lower risk costs under deterministic ones.

Our key contribution has been, in Propositions 7, 8, and 9, to identify environments in which optimally weighted ambiguous contracts dominate all deterministic ones. In each of these environments, optimally weighted ambiguous contracts induce the agent to choose perfectly balanced efforts, and they generate a payoff for the principal arbitrarily close to what he could achieve in the absence of hidden information on the agent's part.

Though our propositions focused on limiting environments to prove analytically that ambiguous schemes can dominate all deterministic ones, our results have more general implications about what characteristics of contracting environments increase the relative attractiveness of ambiguous schemes. Ambiguous schemes are more likely to be preferred when i) the importance to the incentive designer of inducing balanced efforts is high (large δ in our model); ii) even a small amount of ambiguity about the weights in the incentive scheme has a large effect on the agent's chosen degree of effort balance (small λ); and iii) ambiguous schemes provide the agent with a powerful self-insurance motive for balancing efforts (large r). These three factors increase the attractiveness of EAR and EPD in qualitatively the same way. These ambiguous schemes are also more likely to be preferred when iv) the errors in measuring performance on the tasks have small variance and large correlation (small σ^2 and large ρ). In such

environments, EPD provides strong incentives for balanced efforts whatever the relative weights on the performance measures, while under EAR, the cost of manipulating the weights to generate strong incentives for effort balance becomes relatively small.

We emphasize that, because of the agent’s hidden information, ambiguous schemes can dominate deterministic ones even when pay can be based upon measured performance on both tasks. When costs of measurement constrain an incentive designer to base pay on only one performance measure, the attractiveness of ex ante randomization over which task to measure and reward is significantly enhanced relative to the best deterministic scheme rewarding only one task.

Our results have been derived in a static framework, where gaming of an incentive scheme takes the form of excessive focusing of effort on tasks that agents privately find less costly. In dynamic settings where agents’ rewards are based on cumulative performance on each task and agents privately observe interim performance, the type of gaming that Holmström and Milgrom (1987) examined might also arise. There is reason to conjecture that this type of gaming might be more severe under ex post discretion than under ex ante randomization, making the comparison between these schemes more complex. We leave it to future research to explore this question in more detail.

References

- [1] **Baker, George P.**, “Incentive Contracts and Performance Measurement”, *Journal of Political Economy*, Jun. 1992, *100* (3), 598-614.
- [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] **Baliga, Sandeep and Tomas Sjöström**, “Strategic Ambiguity and Arms Proliferation,” *Journal of Political Economy*, Dec. 2008, *116* (6), 1023-1057.
- [4] **Bentham, Jeremy**, *Constitutional Code*, Vol. 1, London: R. Heward, 1830.
- [5] **Bernheim, B. Douglas and Michael D. Whinston**, “Incomplete Contracts and Strategic Ambiguity,” *American Economic Review*, 1998, *88*, 902-932.
- [6] **Bevan, Gwyn and Christopher Hood**, “Targets, Inspections, and Transparency,” *British Medical Journal*, March 2004, *328*, 598.
- [7] **Burgess, Simon, Propper, Carol, Ratto, Marisa and Emma Tominey**, “Incentives in the Public Sector: Evidence from a Government Agency,” *CEPR Discussion Paper 9071*, Jul. 2012.
- [8] **Bond, Philip and Armando Gomes**, “Multitask principal-agent problems: Optimal contracts, fragility, and effort misallocation,” *Journal of Economic Theory*, Jan. 2009, *144* (1), 175-211.
- [9] **Cain, Michael**, “The Moment-Generating Function of the Minimum of Bivariate Normal Random Variables,” *The American Statistician*, May 1994, *48* (2), 124-125.
- [10] **Carrell, Scott E. and James E. West**, “Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors,” *Journal of Political Economy*, Jun. 2010, *118* (3), 409-432.
- [11] **Chetty, Raj, Looney, Adam and Kory Kroft**, “Salience and Taxation: Theory and Evidence,” *American Economic Review*, Sep. 2009, *99* (4), 1145-1177.
- [12] **Dranove, David, Kessler, Daniel, McClellan, Mark and Mark Satterthwaite**, “Is More Information Better? The Effects of ‘Report Cards’ on Health Care Providers,” *Journal of Political Economy*, Jun. 2003, *111* (3), 555-588.
- [13] **Englmaier, Florian, Roider, Andreas and Uwe Sunde**, “The Role of Salience in Performance Schemes: Evidence from a Field Experiment,” *IZA Discussion Paper 6448*, Mar. 2012.
- [14] **Forbes, Silke, Lederman, Mara and Trevor Tombe**, “Quality Disclosure Programs with Thresholds: Misreporting, Gaming and Employee Incentives,” *Working Paper*, 2012.

- [15] **Gjesdal, Froystein**, “Information and Incentives: The Agency Information Problem,” *Review of Economic Studies*, Jul. 1982, *49* (3), 373-390.
- [16] **Gravelle, Hugh, Sutton, Matt and Ada Ma**, “Doctor Behaviour under a Pay for Performance Contract: Treating, Cheating and Case Finding?” *Economic Journal*, Feb. 2010, *120* (1), F129-F156.
- [17] **Grossman, Sanford J. and Oliver D. Hart**, “An Analysis of the Principal-Agent Problem,” *Econometrica*, Jan. 1983, *51* (1), 7-45.
- [18] **Healy, Andrew and Jennifer Pate**, “Can Teams Help to Close the Gender Competition Gap?” *Economic Journal*, Sep. 2011, *121* (6), 1192-1204.
- [19] **Holmström, Bengt and Paul Milgrom**, “Aggregation and Linearity in the Provision of Intertemporal Incentives,” *Econometrica*, Mar. 1987, *55* (2), 303-328.
- [20] — and —, “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design,” *Journal of Law, Economics, and Organization*, Jan. 1991, *7* (Special Issue: Papers from the Conference on the New Science of Organization), 24-52.
- [21] **Jehiel, Philippe**, “On Transparency in Organizations,” *Working Paper*, 2011.
- [22] **Jehiel, Philippe and Andrew F. Newman**, “Loopholes: Social Learning and the Evolution of Contract Form,” *Working Paper*, 2011.
- [23] **Lazear, Edward P.**, “Speeding, Terrorism, and Teaching to the Test,” *Quarterly Journal of Economics*, Aug. 2006, *121* (3), 1029-1061.
- [24] **MacDonald, Glenn and Leslie M. Marx**, “Adverse Specialization,” *Journal of Political Economy*, Aug. 2001, *109* (4), 864-899.
- [25] **MacLeod, W. Bentley and James M. Malcomson**, “Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment,” *Econometrica*, Mar. 1989, *57* (2), 447-480.
- [26] **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.
- [27] **Osler, M.**, “The Case Against Transparency in Law School Rankings Methodology,” *Law School Innovation*, May 4, 2010.
- [28] **Oyer, Paul**, “Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality,” *Quarterly Journal of Economics*, Feb. 1998, *113* (1), 149-185.
- [29] **Pestieau, Pierre, Possen, Uri M. and Steven M. Slutsky**, “The Value of Explicit Randomization in the Tax Code,” *Journal of Public Economics*, Jan. 1998, *67*(1), 87-103.
- [30] **Prendergast, Canice**, “The Provision of Incentives in Firms,” *Journal of Economic Literature*, Mar. 1999, *37* (1), 7-63.
- [31] **Rahman, David**, “But Who Will Monitor the Monitor?” *American Economic Review*, Oct. 2012, *102* (6), 2767-2797.
- [32] **Scott, Robert E. and George G. Triantis**, “Anticipating Litigation in Contract Design,” *Yale Law Journal*, Jan. 2006, *115* (4), 814-879.
- [33] **Stiglitz, Joseph**, “Self-selection and Pareto Efficient Taxation,” *Journal of Public Economics*, Mar. 1982, *17* (2), 213-240.
- [34] **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. Under the SD contract $w = \alpha + \beta x_1 + \beta x_2$, the agent is willing to set $e_1 = e_2 = \frac{\beta}{2}$, and the agent’s certainty equivalent is

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

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

$$\Pi^{SD}(\beta; \lambda = 1) = \frac{\beta}{2} \left(1 + \frac{1}{\delta}\right) - \frac{\beta^2}{2} - r\sigma^2\beta^2(1 + \rho). \quad (23)$$

With β chosen optimally, the resulting maximized payoff is $\Pi^{SD}(\lambda = 1) = \frac{(\delta+1)^2}{8\delta^2[1+2r\sigma^2(1+\rho)]}$. Under the OT contract $w = \alpha + \beta x_1 - \rho\beta x_2$, the agent sets $e_1 = \beta$ and $e_2 = 0$. With α chosen optimally, the principal's expected payoff as a function of β is

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

and the optimal choice of β yields payoff

$$\Pi^{OT} = \frac{1}{2\delta^2[1 + r\sigma^2(1 - \rho^2)]}. \quad (24)$$

Comparison of the maximized payoff expressions shows that there is a critical value of δ ,

$$\delta^1(r\sigma^2, \rho) \equiv 2 \left[\frac{1 + 2r\sigma^2(1 + \rho)}{1 + r\sigma^2(1 - \rho^2)} \right]^{\frac{1}{2}} - 1, \quad (25)$$

above (below) which $\Pi^{SD}(\lambda = 1) > (<) \Pi^{OT}$. It is simple to show that $\delta^1(r\sigma^2, \rho)$ is greater than 1 and increasing in each of its arguments. ■

Proof of Lemma 2. Consider first the pair of contracts (C_1^{bal}, C_2^{bal}) . C_i^{bal} induces agent i to choose $e_i = e_j = \frac{\beta}{1+\lambda}$, yielding each type i a certainty equivalent of

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

The principal will set α to satisfy each type's participation constraint with equality, and his expected payoff from each type, as a function of β , will be

$$\Pi^{bal}(\beta) = \frac{\beta}{1+\lambda} \left(1 + \frac{1}{\delta}\right) - \frac{\beta^2}{2} - \frac{1}{2}r\sigma^2\beta^2(1 + 2\rho\lambda + \lambda^2). \quad (26)$$

With β chosen optimally, the resulting maximized payoff is

$$\Pi^{bal} = \frac{(\delta + 1)^2}{2\delta^2(1 + \lambda)^2 [1 + r\sigma^2(1 + 2\rho\lambda + \lambda^2)]}. \quad (27)$$

This payoff is continuous as $\lambda \rightarrow 1$ and its limiting value is $\Pi^{SD}(\lambda = 1) = \frac{(\delta+1)^2}{8\delta^2[1+2r\sigma^2(1+\rho)]}$.

Now consider the pair of contracts (C_1^{OT}, C_2^{OT}) . C_i^{OT} induces type i to choose $e_i = \beta$ and $e_j = 0$. The principal's payoff from each type of agent, when α and β are chosen optimally, is given by expression (24) derived previously for the OT contract.

Comparison of the payoff expressions (27) and (24) shows that there is a critical value of δ ,

$$\delta^{NHI}(\lambda, r\sigma^2, \rho) \equiv (\lambda + 1) \left[\frac{1 + r\sigma^2(1 + 2\rho\lambda + \lambda^2)}{1 + r\sigma^2(1 - \rho^2)} \right]^{\frac{1}{2}} - 1, \quad (28)$$

above (below) which $\Pi^{bal} > (<) \Pi^{OT}$. It is easy to show that $\delta^{NHI}(\lambda, r\sigma^2, \rho)$ is increasing in each of its arguments and that, as λ approaches 1, δ^{NHI} approaches δ^1 . ■

Proof of Lemma 3. Inducing both types of agent to choose perfectly balanced efforts would necessitate both

i) offering a menu of the form

$$C_1 : w_1 = \alpha_1 + \beta_1 x_1 + \lambda \beta_1 x_2 \quad \text{and} \quad C_2 : w_2 = \alpha_2 + \beta_2 x_2 + \lambda \beta_2 x_1,$$

and ii) inducing agent i to choose C_i from this menu. Let $ACE_i(C_j)$ denote the certainty equivalent achieved by agent i from choosing contract C_j . For agent 1 to be willing to choose C_1 requires $ACE_1(C_1) \geq ACE_1(C_2)$, and the analogous self-selection constraint for agent 2 is $ACE_2(C_2) \geq ACE_2(C_1)$. Now for all $\lambda > 1$, $ACE_2(C_1) > ACE_1(C_1)$, since agent 1's certainty equivalent from contract C_1 equals that which he would obtain from focusing all his effort on task 1 (which is one of his optimal effort allocations), whereas agent 2's certainty equivalent from C_1 equals that which he would obtain from focusing all his effort on task 2 (which is his unique optimal effort choice), and task 2 is more highly rewarded than task 1 in contract C_1 . Similarly, for all $\lambda > 1$, $ACE_1(C_2) > ACE_2(C_2)$. If $ACE_1(C_1) \geq ACE_2(C_2)$, then $ACE_2(C_1) > ACE_1(C_1)$ implies that $ACE_2(C_1) > ACE_2(C_2)$, so the self-selection constraint for agent 2 would be violated. If, instead, $ACE_1(C_1) < ACE_2(C_2)$, then $ACE_1(C_2) > ACE_2(C_2)$ implies that $ACE_1(C_1) < ACE_1(C_2)$, so the self-selection constraint for agent 1 would be violated. Therefore, there is no way to choose $(\alpha_1, \beta_1, \alpha_2, \beta_2)$ so that the menu above induces both types of privately-informed agent to choose the contract that would make each willing to choose perfectly balanced efforts. Hence perfectly balanced efforts from both types of agents cannot be achieved.

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

Proof of Proposition 1. Part i) is proved in the text immediately following the statement of the proposition. For part ii), consider first the SDM, consisting of the contract pair

$$C_1^{SDM} : w_1 = \alpha + \beta x_1 - \rho \beta x_2 \quad \text{and} \quad C_2^{SDM} : w_2 = \alpha + \beta x_2 - \rho \beta x_1.$$

Agent i strictly prefers contract C_i to contract C_j and, having chosen C_i , will then set $e_i = \beta$ and $e_j = 0$. This SDM 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}) . Therefore, the principal's maximized payoff from a SDM, Π^{SDM} , is given by the expression in (24). (Compared to a SDM, a SD contract would, for all $\lambda > 1$, also induce fully focused efforts from both agent types but would impose a larger risk premium and hence generate a lower payoff for the principal.)

Now consider the ADM, consisting of the contract pair

$$C_1^{ADM} : w_1 = \alpha_1 + \beta_1 x_1 - \rho \beta_1 x_2 \quad \text{and} \quad C_2^{ADM} : w_2 = \alpha_2 + \beta_2 x_2 + \lambda \beta_2 x_1.$$

In this menu, C_i is the contract intended for agent i . If agent 2 chooses C_2 , he would be willing to choose $e_1 = e_2 = \frac{\beta_2}{1+\lambda}$. If, instead, he chooses C_1 , he would set $e_1 = \frac{\beta_1}{\lambda^2}$ and $e_2 = 0$. If agent 1 chooses C_1 , he would set $e_1 = \beta_1$ and $e_2 = 0$, whereas if he chooses C_2 , he would set $e_1 = \lambda \beta_2$ and $e_2 = 0$.

The certainty equivalents that each of C_1 and C_2 offers to each type of agent are:

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

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

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

Since for all $\lambda > 1$ we have $ACE_1(C_2) > ACE_2(C_2)$, agent 1's participation constraint will not bind, and hence

agent 1 earns an “information rent”.

For the two self-selection constraints to be satisfied simultaneously, it is necessary that $\beta_1 \geq \lambda\beta_2$. 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

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

There exists a $\hat{\delta}$ such that the constraint $\beta_1 \geq \lambda\beta_2$ will be binding at the optimum if and only if $\delta \geq \hat{\delta}$. If $\delta < \hat{\delta}$, then the principal’s maximized payoff from this “unconstrained” ADM (ADMU) is

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

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

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

It remains to determine whether an ADM (unconstrained or constrained) or a SDM is optimal. It can be checked that the crucial comparison is between Π^{SDM} and Π^{ADMU} and furthermore that if

$$\delta < \delta^{HI}(\lambda, r\sigma^2, \rho) \equiv (\lambda + 1) \sqrt{\frac{\lambda^2 + r\sigma^2(\lambda^2 + 2\rho\lambda + 1)}{1 + r\sigma^2(1 - \rho^2)}} - 1,$$

then the best SDM dominates the best ADM, whereas if $\delta > \delta^{HI}(\lambda, r\sigma^2, \rho)$, then the best ADM dominates the best SDM. $\delta^{HI} < \hat{\delta}$. This proves part (ii). The claims in parts (iii) and (iv) are easily shown. ■

Proof of Proposition 2.

Proof of Parts (i) and (ii): For each type of agent, let \bar{e} (respectively, \underline{e}) denote effort on his less costly (respectively, more costly) task, and define \bar{x} and \underline{x} analogously. Under ex ante randomization, with probability $\frac{1}{2}$, $w = \alpha + \beta\bar{x} + k\beta\underline{x}$, in which case we let \overline{EU} denote an agent’s expected utility, and with probability $\frac{1}{2}$, $w = \alpha + \beta\underline{x} + k\beta\bar{x}$, in which case we denote expected utility by \underline{EU} .

Recall that $k \in (-1, 1)$. Each agent’s unconditional expected utility under EAR is

$$\begin{aligned} \frac{1}{2}\overline{EU} + \frac{1}{2}\underline{EU} &= -\frac{1}{2}E \exp \left\{ -r \left[\alpha + \beta\bar{x} + k\beta\underline{x} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 \right] \right\} - \frac{1}{2}E \exp \left\{ -r \left[\alpha + \beta\underline{x} + k\beta\bar{x} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 \right] \right\} \\ &= -\frac{1}{2} \exp \left\{ -r \left[\alpha + \beta\bar{e} + k\beta\underline{e} - \frac{r}{2}\sigma^2\beta^2(1 + 2\rho k + k^2) - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 \right] \right\} \\ &\quad - \frac{1}{2} \exp \left\{ -r \left[\alpha + \beta\underline{e} + k\beta\bar{e} - \frac{r}{2}\sigma^2\beta^2(1 + 2\rho k + k^2) - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 \right] \right\} \quad (29) \end{aligned}$$

Hence the first-order conditions for interior solutions for \bar{e} and \underline{e} , respectively, are

$$\begin{aligned} \frac{1}{2} [\beta - (\bar{e} + \lambda\underline{e})] \overline{EU} + \frac{1}{2} [k\beta - (\bar{e} + \lambda\underline{e})] \underline{EU} &= 0 \\ \frac{1}{2} [k\beta - \lambda(\bar{e} + \lambda\underline{e})] \overline{EU} + \frac{1}{2} [\beta - \lambda(\bar{e} + \lambda\underline{e})] \underline{EU} &= 0. \end{aligned}$$

These first-order conditions can be rewritten as

$$\beta\overline{EU} + k\beta\underline{EU} = (\bar{e} + \lambda e)(\overline{EU} + \underline{EU}) \quad (30)$$

$$k\beta\overline{EU} + \beta\underline{EU} = \lambda(\bar{e} + \lambda e)(\overline{EU} + \underline{EU}). \quad (31)$$

Equations (30) and (31) in turn imply

$$\overline{EU} + k\underline{EU} = \frac{k}{\lambda}\overline{EU} + \frac{1}{\lambda}\underline{EU}.$$

If $k \in [\frac{1}{\lambda}, 1)$, then the left-hand side of this equation strictly exceeds the right-hand side, so in this case interior solutions for efforts cannot exist. This proves Part (i).

Adding the first-order conditions (30) and (31) and rearranging yields equation (1). Using (1) to substitute for aggregate effort $(\bar{e} + \lambda e)$ in (30) yields, after a little algebra, $(\lambda - k)\overline{EU} + (k\lambda - 1)\underline{EU} = 0$, which simplifies to equation (2).

Proof of Part (iii): Solving (2) for $\bar{e} - e$ yields $\bar{e} - e = [\ln(\frac{\lambda-k}{1-k\lambda})]/[r\beta(1-k)]$. For $k \in (-1, \frac{1}{\lambda})$ and $\lambda > 1$, therefore, $(\bar{e} - e)$ is greater than 0, increasing in λ and k , and decreasing in r . $(\bar{e} - e) \rightarrow 0$ as $\lambda \rightarrow 1$, $k \rightarrow -1^+$, or $r \rightarrow \infty$.

Proof of Part (iv): Using (1) and (2) to substitute into (29), and then simplifying, allows us to express each type of agent's expected utility under EAR as

$$\frac{1}{2}\overline{EU} + \frac{1}{2}\underline{EU} = -\exp\left\{-r\left[\alpha + \beta(\bar{e} + ke) - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2\beta^2(1+2\rho k+k^2) - \frac{1}{r}\ln\left(\frac{1+\frac{\lambda-k}{1-k\lambda}}{2}\right)\right]\right\}.$$

Since both types receive the same expected utility, it is optimal for the principal to set α to ensure that their participation constraints bind. With α set in this way (so that the whole expression in square brackets above is equal to 0), the principal's expected payoff, for given (β, k) , can be simplified to equation (3) as follows:

$$\begin{aligned} \Pi^{EAR}(\beta, k) &= e + \frac{1}{\delta}\bar{e} - \alpha - \frac{1}{2}\beta(\bar{e} + ke) - \frac{1}{2}\beta(e + k\bar{e}) \\ &= e + \frac{1}{\delta}\bar{e} + \frac{1}{2}\beta(1-k)(\bar{e} - e) - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2\beta^2(1+2\rho k+k^2) - \frac{1}{r}\ln\left(\frac{1+\frac{\lambda-k}{1-k\lambda}}{2}\right) \\ &= e + \frac{1}{\delta}\bar{e} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2\beta^2(1+2\rho k+k^2) - \frac{1}{2r}\ln\left[\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right], \end{aligned}$$

where the final line uses (2). ■

Proof of Proposition 3. Define $B \equiv \beta(1+k)$ and note, from (1), that aggregate effort $\bar{e} + \lambda e$ is proportional to B . Using (1), (2), and $\beta = \frac{B}{1+k}$ to substitute into (3) yields (5) and (6) in the text. To prove the claims regarding the effect of varying δ , r (with $r\sigma^2$ fixed), or $\sigma^2(1-\rho)$ on the optimal level of k , we use (5) and (6) to examine the sign of the cross-partial derivative of $\Pi^{EAR}(B, k)$ with respect to k and the relevant parameter, holding B and hence aggregate effort fixed.

Part (i): Only the second term on the right-hand side of (6) generates a non-zero value of $\frac{\partial^2\Pi}{\partial\delta\partial k}$. As long as $\delta > \lambda$, $\frac{\partial^2\Pi}{\partial\delta\partial k} < 0$, so the optimal k decreases as δ increases.

Part (ii): With $r\sigma^2$ held fixed, only the second term on the right-hand side of (6) and fourth term in (5) vary as r increases. Examining these terms shows that $\frac{\partial^2\Pi}{\partial r\partial k} > 0$, so as r decreases (holding $r\sigma^2$ fixed), the optimal k decreases.

Part (iii): $\frac{\partial\Pi}{\partial k}$ depends on σ^2 and ρ only via the third term in (5), and $\frac{\partial\Pi}{\partial k}$ is increasing in $\sigma^2(1-\rho)$, so the optimal k decreases as $\sigma^2(1-\rho)$ decreases.

Finally, $\frac{\partial^2\Pi}{\partial B\partial k} > 0$, so as the B to be induced decreases, the optimal k decreases. ■

Proof of Proposition 4. The proof parallels the steps of the proof of Proposition 2.

Proof of Parts (i) and (ii): Define \bar{e} , e , \bar{x} , and \underline{x} as in the proof of Proposition 2. Under ex post discretion, since each type of agent anticipates that he will receive the wage $w = \alpha + \beta \min\{\bar{x} + k\underline{x}, \underline{x} + k\bar{x}\}$, each type's

expected utility is

$$\begin{aligned} & - E \exp \left\{ -r \left[\alpha + \beta \min \{ \bar{x} + k\underline{x}, \underline{x} + k\bar{x} \} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 \right] \right\} \\ & = - \exp \left\{ -r \left[\alpha - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 \right] \right\} E \{ \exp [-r\beta \min \{ \bar{x} + k\underline{x}, \underline{x} + k\bar{x} \}] \}. \end{aligned} \quad (32)$$

The random variables $(\bar{x} + k\underline{x})$ and $(\underline{x} + k\bar{x})$ have a bivariate normal distribution, with means $\bar{e} + k\underline{e}$ and $\underline{e} + k\bar{e}$, respectively. Their common variance and correlation coefficient are, respectively,

$$(\sigma^k)^2 \equiv \text{var} (x_1 + kx_2) = \sigma^2 (1 + 2\rho k + k^2) \quad \text{and} \quad \rho^k \equiv \text{corr} (x_1 + kx_2, x_2 + kx_1) = \frac{\rho + 2k + \rho k^2}{1 + 2\rho k + k^2}. \quad (33)$$

Denote by $\Phi(\cdot)$ and $\phi(\cdot)$ the c.d.f. and p.d.f., respectively, of a standard normal random variable. Define

$$\theta \equiv \sigma [2(1 - \rho)]^{\frac{1}{2}} \quad \text{and} \quad \theta^k \equiv \sigma^k [2(1 - \rho^k)]^{\frac{1}{2}} = \theta(1 - k). \quad (34)$$

Cain (1994) derived the moment-generating function, $m(t)$, for the minimum of bivariate normal random variables. Using his formula and the definitions of $\Phi(-)$ and $\Phi(+)$ in the statement of Proposition 4, we have

$$\begin{aligned} m(-r\beta) & \equiv E \{ \exp [-r\beta \min \{ \bar{x} + k\underline{x}, \underline{x} + k\bar{x} \}] \} \\ & \quad + \exp \left[-r\beta (\underline{e} + k\bar{e}) + \frac{1}{2}r^2\beta^2 (\sigma^k)^2 \right] \Phi \left(\frac{(\bar{e} - \underline{e})(1 - k) + r\beta (\sigma^k)^2 (1 - \rho^k)}{\theta^k} \right) \\ & = \exp \left[-r\beta (\bar{e} + k\underline{e}) + \frac{1}{2}r^2\beta^2 (\sigma^k)^2 \right] \Phi \left(\frac{-(\bar{e} - \underline{e})}{\theta} + \frac{r\beta\theta(1 - k)}{2} \right) \\ & \quad + \exp \left[-r\beta (\underline{e} + k\bar{e}) + \frac{1}{2}r^2\beta^2 (\sigma^k)^2 \right] \Phi \left(\frac{(\bar{e} - \underline{e})}{\theta} + \frac{r\beta\theta(1 - k)}{2} \right) \\ & = \exp \left[-r\beta (\bar{e} + k\underline{e}) + \frac{1}{2}r^2\beta^2 (\sigma^k)^2 \right] \Phi(-) + \exp \left[-r\beta (\underline{e} + k\bar{e}) + \frac{1}{2}r^2\beta^2 (\sigma^k)^2 \right] \Phi(+). \end{aligned} \quad (35)$$

Using (35) in (32), we can derive and simplify the first-order conditions for interior optimal values of \bar{e} and \underline{e} , respectively:

$$(\bar{e} + \lambda\underline{e})m(-r\beta) = \beta \exp \left[-r\beta (\bar{e} + k\underline{e}) + \frac{1}{2}r^2\beta^2 (\sigma^k)^2 \right] \Phi(-) + k\beta \exp \left[-r\beta (\underline{e} + k\bar{e}) + \frac{1}{2}r^2\beta^2 (\sigma^k)^2 \right] \Phi(+). \quad (36)$$

$$\lambda(\bar{e} + \lambda\underline{e})m(-r\beta) = k\beta \exp \left[-r\beta (\bar{e} + k\underline{e}) + \frac{1}{2}r^2\beta^2 (\sigma^k)^2 \right] \Phi(-) + \beta \exp \left[-r\beta (\underline{e} + k\bar{e}) + \frac{1}{2}r^2\beta^2 (\sigma^k)^2 \right] \Phi(+). \quad (37)$$

Equations (36) and (37) imply

$$\exp[-r\beta (\bar{e} + k\underline{e})]\Phi(-) + k \exp[-r\beta (\underline{e} + k\bar{e})]\Phi(+)= \frac{k}{\lambda} \exp[-r\beta (\bar{e} + k\underline{e})]\Phi(-) + \frac{1}{\lambda} \exp[-r\beta (\underline{e} + k\bar{e})]\Phi(+).$$

If $k \in [\frac{1}{\lambda}, 1)$, then the left-hand side of this equation strictly exceeds the right-hand side, so in this case interior solutions for efforts cannot exist. This proves Part (i).

Adding the first-order conditions (36) and (37) and rearranging yields equation (7). Using (7) to substitute for aggregate effort $(\bar{e} + \lambda\underline{e})$ in (36) yields, after simplification, equation (8).

Proof of Part (iii): Define $d \equiv \bar{e} - \underline{e}$ and rewrite (8) in terms of d as

$$\exp [r\beta (1 - k) d] \frac{\Phi \left(\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2} \right)}{\Phi \left(-\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2} \right)} = \frac{\lambda - k}{1 - k\lambda}. \quad (38)$$

The left-hand side of (38) is increasing in d and equals 1 at $d = 0$. The right-hand side is increasing in λ and $\rightarrow 1$ as $\lambda \rightarrow 1$, so d is increasing in λ and $d \rightarrow 0$ as $\lambda \rightarrow 1$. Differentiating (38) implicitly w.r.t. k yields

$$\begin{aligned} \frac{\partial}{\partial k} \left(\frac{\lambda - k}{1 - k\lambda} \right) &= \exp[r\beta(1-k)d] \frac{\Phi(+)}{\Phi(-)} \left[-r\beta d + r\beta(1-k) \frac{\partial d}{\partial k} \right] \\ &\quad + \frac{\exp[r\beta(1-k)d]}{(\Phi(-))^2} \left[\phi(+)\Phi(-) \left(\frac{1}{\theta} \frac{\partial d}{\partial k} - \frac{r\beta\theta}{2} \right) - \phi(-)\Phi(+) \left(-\frac{1}{\theta} \frac{\partial d}{\partial k} - \frac{r\beta\theta}{2} \right) \right], \end{aligned}$$

where $\phi(+)$ and $\phi(-)$ are defined analogously to $\Phi(+)$ and $\Phi(-)$. From this, we can conclude that

$$\frac{\partial d}{\partial k} \stackrel{\text{sgn}}{=} (\Phi(-))^2 \frac{\partial}{\partial k} \left(\frac{\lambda - k}{1 - k\lambda} \right) + r\beta \exp(r\beta(1-k)d) \left[d\Phi(+)\Phi(-) + \frac{\theta}{2} (\phi(+)\Phi(-) - \phi(-)\Phi(+)) \right].$$

Hence to show that $\frac{\partial d}{\partial k} \geq 0$, it is sufficient to show that the term in square brackets on the right-hand side above is positive, since $\frac{\partial}{\partial k} \left(\frac{\lambda - k}{1 - k\lambda} \right) \geq 0$. Now define

$$y \equiv \frac{d}{\theta} \quad \text{and} \quad t \equiv \frac{r\beta\theta(1-k)}{2}. \quad (39)$$

With these definitions, the term in square brackets above has the sign of

$$\frac{2d}{\theta} + \frac{\phi(+)}{\Phi(+)} - \frac{\phi(-)}{\Phi(-)} = y + t + \frac{\phi(y+t)}{\Phi(y+t)} - \left(-y + t + \frac{\phi(-y+t)}{\Phi(-y+t)} \right) = j(y+t) - j(-y+t), \quad (40)$$

where $j(z) \equiv z + \frac{\phi(z)}{\Phi(z)}$. Since $y \geq 0$ and $j(\cdot)$ is increasing, $j(y+t) - j(-y+t) \geq 0$. Hence d is increasing in k . As $k \rightarrow -1^+$, the right-hand side of (38) approaches 1, so d approaches 0.

Differentiating (38) implicitly with respect to r and rearranging shows that $\frac{\partial d}{\partial r}$ has the opposite sign to the expressions in (40), so d is decreasing in r . As $r \rightarrow \infty$, the ratio $\frac{\Phi(+)}{\Phi(-)}$ on the left-hand side of (38) approaches 1, so for the left-hand side as a whole to remain finite requires $d \rightarrow 0$. Hence as $r \rightarrow \infty$, $d \rightarrow 0$.

To prove that d is increasing in $\sigma^2(1-\rho) = \frac{\theta^2}{2}$, differentiate (38) implicitly with respect to θ . This yields

$$\frac{\partial d}{\partial \theta} \stackrel{\text{sgn}}{=} \frac{d}{\theta^2} [\phi(+)\Phi(-) + \phi(-)\Phi(+)] - \frac{r\beta(1-k)}{2} [\phi(+)\Phi(-) - \phi(-)\Phi(+)].$$

Since $\frac{\phi(\cdot)}{\Phi(\cdot)}$ is decreasing, $\phi(+)\Phi(-) - \phi(-)\Phi(+) < 0$, so d is increasing in θ and hence in $\sigma^2(1-\rho)$. As $\sigma^2(1-\rho)$ and hence θ goes to 0, $d \rightarrow 0$, since otherwise the left-hand side of (38) becomes infinite.

Proof of Part (iv): Using (7), (8), and (35) to substitute into (32) allows us to express each type of agent's expected utility under EPD as

$$\begin{aligned} & - \exp \left\{ -r \left[\alpha - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2} r \beta^2 (\sigma^k)^2 \right] \right\} (\exp[-r\beta(\bar{e} + k\underline{e})]\Phi(-) + \exp[-r\beta(\underline{e} + k\bar{e})]\Phi(+)) \\ & = - \exp \left\{ -r \left[\alpha - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2} r \beta^2 (\sigma^k)^2 \right] \right\} (\exp[-r\beta(\underline{e} + k\bar{e})] \{ \exp[-r\beta(1-k)d]\Phi(-) + \Phi(+)\}). \end{aligned}$$

Hence each type of agent's certainty equivalent is

$$\alpha + \beta(\underline{e} + k\bar{e}) - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2} r \beta^2 (\sigma^k)^2 - \frac{1}{r} \ln [\exp\{-r\beta(1-k)d\}\Phi(-) + \Phi(+)].$$

Using Cain's formula for the expectation of the minimum of bivariate normal random variables, as well as the

definitions of d , θ^k , and θ , the principal's payoff under EPD can be expressed as

$$\begin{aligned} & \underline{e} + \frac{1}{\delta} \bar{e} - \alpha - \beta E \min\{x_1 + kx_2, x_2 + kx_1\} \\ &= \underline{e} + \frac{1}{\delta} \bar{e} - \alpha - \beta \left[(\bar{e} + k\underline{e}) \Phi \left(\frac{-(1-k)d}{\theta^k} \right) + (\underline{e} + k\bar{e}) \Phi \left(\frac{(1-k)d}{\theta^k} \right) - \theta^k \phi \left(\frac{(1-k)d}{\theta^k} \right) \right] \\ &= \underline{e} + \frac{1}{\delta} \bar{e} - \alpha - \beta \left[(\bar{e} + k\underline{e}) \Phi \left(\frac{-d}{\theta} \right) + (\underline{e} + k\bar{e}) \Phi \left(\frac{d}{\theta} \right) - \theta(1-k) \phi \left(\frac{d}{\theta} \right) \right]. \end{aligned}$$

The principal will set α so each agent's certainty equivalent is 0, yielding a payoff for the principal of:

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

which is the payoff expression given in (9).

Proof of Part (v): The left-hand sides of (2) and (8) are both increasing in $d \equiv (\bar{e} - \underline{e})$, and since $\Phi(+)$ > $\Phi(-)$ whenever $d > 0$, the left-hand side of (8) is strictly greater than the left-hand side of (2) for all $d > 0$. Since $\lambda > 1$ implies that $d^{EAR} > 0$ and $d^{EPR} > 0$, it follows that for all $\lambda > 1$, $d^{EPD} < d^{EAR}$. ■

Proof of Proposition 5. Equations (3) and (9) give the principal's payoff from interior effort choices by the agents under EAR and EPD, respectively, for given (β, k) . The proof proceeds in three steps:

Step 1:

$$\Pi^{EPD}(\beta, k) - \left[\underline{e}^{EPD} + \frac{1}{\delta} \bar{e}^{EPD} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2} r(\sigma)^2 \beta^2 (1+2\rho k + k^2) \right] \geq 0, \quad (41)$$

(41) says that for any (β, k) , EPD imposes lower risk costs than would the deterministic contract $w = \alpha + \beta x_1 + k\beta x_2$. To prove (41), we use (9) to express the left-hand side as

$$\begin{aligned} & \frac{1}{r} \left\{ -r\beta d(1-k) \Phi \left(\frac{-d}{\theta} \right) + r\beta\theta(1-k) \phi \left(\frac{d}{\theta} \right) \right. \\ & \quad \left. - \ln \left[\exp\{-r\beta(1-k)d\} \Phi \left(\frac{-d}{\theta} + \frac{r\beta\theta(1-k)}{2} \right) + \Phi \left(\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2} \right) \right] \right\}. \quad (42) \end{aligned}$$

The terms in curly brackets in (42) can be rewritten as

$$h(y, t) \equiv -2ty\Phi(-y) + 2t\phi(-y) - \ln [\exp\{-2ty\} \Phi(-y+t) + \Phi(y+t)], \quad (43)$$

where we have used the definitions of y and t in (39) and the fact that $\phi(y) = \phi(-y)$. We now show that $h(y, t) \geq 0$ for all $y \geq 0, t \geq 0$. First observe that $h(0, t) = 2t\phi(0) - \ln[2\Phi(t)]$ and $h(0, 0) = 0$. Also,

$$\frac{\partial h(0, t)}{\partial t} = 2\phi(0) - \frac{\phi(t)}{\Phi(t)},$$

and since $\frac{\partial h(0, t)}{\partial t} = 0$ at $t = 0$ and $\frac{\phi(t)}{\Phi(t)}$ is decreasing in t , it follows that $h(0, t) \geq 0$ for all $t \geq 0$. Furthermore, as $y \rightarrow \infty$, $h(y, t) \rightarrow 0$ for all $t \geq 0$. Thus, to show that for all $y \geq 0$ and $t \geq 0$, $h(y, t) \geq 0$, it is sufficient to show that $\frac{\partial h(y, t)}{\partial y} \leq 0$. Now

$$\frac{\partial h(y, t)}{\partial y} \stackrel{\text{sgn}}{=} -\Phi(-y) + \frac{\Phi(-y+t)}{\exp\{2ty\} \Phi(y+t) + \Phi(-y+t)}. \quad (44)$$

Define $q(y, t)$ to equal the right-hand side of (44). $q(y, 0) = 0$, so showing that $q(y, t)$ is decreasing in t for all

$t \geq 0$ will imply that $\frac{\partial h(y,t)}{\partial y} \leq 0$.

$$\frac{\partial q(y,t)}{\partial t} \stackrel{\text{sgn}}{=} -y + t + \frac{\phi(-y+t)}{\Phi(-y+t)} - \left(y + t + \frac{\phi(y+t)}{\Phi(y+t)} \right) \leq 0,$$

since $j(z) = z + \frac{\phi(z)}{\Phi(z)}$ is increasing (see (40)). Hence for all $y \geq 0, t \geq 0$, $\frac{\partial h(y,t)}{\partial y} \leq 0$, and so $h(y,t) \geq 0$.

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

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

This step follows from the facts that aggregate effort $\bar{e} + \lambda \underline{e}$ is equal under EPD and EAR (as shown by (1) and (7)) and that the gap in efforts, $\bar{e} - \underline{e}$, is smaller under EPD than EAR.

Step 3:

$$\left[\underline{e}^{EAR} + \frac{1}{\delta} \bar{e}^{EAR} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2} r(\sigma^k)^2 \beta^2 \right] - \Pi^{EAR}(\beta, k) = \frac{1}{2r} \ln \left[\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)} \right] \geq 0.$$

This step follows from (3), $\lambda \geq 1$, and the fact that $k < \frac{1}{\lambda}$ is a necessary condition for EAR and EPD to induce interior solutions for efforts. ■

Proof of Proposition 6.

Proof of Part (i): 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 strictly dominated in terms of payoffs by a suitably designed symmetric deterministic (SD) scheme.

For $\lambda = 1$, aggregate effort under EPD is $\bar{e}^{EPD} + \lambda \underline{e}^{EPD} = \frac{\beta(1+k)}{2}$, and $\bar{e}^{EPD} = \underline{e}^{EPD} = \frac{\beta(1+k)}{4}$. Hence, for $\lambda = 1$, equation (9) simplifies to

$$\Pi^{EPD}(\beta, k) = \frac{\delta+1}{\delta} \frac{\beta(1+k)}{4} - \frac{1}{8} \beta^2(1+k)^2 - \frac{1}{2} r(\sigma^k)^2 \beta^2 - \frac{1}{r} \left\{ \ln \left[2\Phi \left(\frac{r\beta\theta(1-k)}{2} \right) \right] - r\beta\theta(1-k)\phi(0) \right\}. \quad (45)$$

Consider now a SD scheme with coefficient β^{SD} chosen to induce the same level of aggregate effort as under EPD for the given values of β and k : $\beta^{SD} = \frac{\beta(1+k)}{2}$. Since $\lambda = 1$, $\bar{e}^{SD} = \underline{e}^{SD} = \frac{\beta(1+k)}{4}$, so SD also induces exactly the same effort levels on each task as EPD. The principal's payoff under the SD scheme is

$$\Pi^{SD}(\beta^{SD}) = \frac{\delta+1}{\delta} \frac{\beta^{SD}}{2} - \frac{1}{2} (\beta^{SD})^2 - r\sigma^2 (\beta^{SD})^2 (1+\rho) = \frac{\delta+1}{\delta} \frac{\beta(1+k)}{4} - \frac{1}{8} \beta^2(1+k)^2 - \frac{1}{4} r\sigma^2 \beta^2(1+k)^2(1+\rho). \quad (46)$$

Using (45) and (46) and the definition of $(\sigma^k)^2$ in (33), we have

$$\begin{aligned} \Pi^{SD}(\beta^{SD}) - \Pi^{EPD}(\beta, k) &= \frac{r\beta^2}{2} \left[(\sigma^k)^2 - \frac{\sigma^2(1+k)^2(1+\rho)}{2} \right] + \frac{1}{r} \left\{ \ln \left[2\Phi \left(\frac{r\beta\theta(1-k)}{2} \right) \right] - r\beta\theta(1-k)\phi(0) \right\} \\ &= \frac{1}{4} r\sigma^2 \beta^2 (1-\rho)(1-k)^2 + \frac{1}{r} \left\{ \ln \left[2\Phi \left(\frac{r\beta\theta(1-k)}{2} \right) \right] - r\beta\theta(1-k)\phi(0) \right\}. \end{aligned} \quad (47)$$

(47) has the sign of $g(t) \equiv \frac{t^2}{2} + \ln[2\Phi(t)] - 2t\phi(0)$, where t is defined in (39). We have $g(0) = 0$. Also,

$$g'(t) = -\sqrt{\frac{2}{\pi}} + \frac{t\Phi(t) + \phi(t)}{\Phi(t)} \text{ and } g'(0) = 0; \quad g''(t) = \frac{[\Phi(t)]^2 - t\phi(t)\Phi(t) - [\phi(t)]^2}{[\Phi(t)]^2} \text{ and } g''(0) \stackrel{\text{sgn}}{=} \frac{1}{4} - \frac{1}{2\pi} > 0$$

Finally, the derivative of the numerator of $g''(t)$ can be shown to be strictly positive for all $t > 0$. Therefore, for all $t > 0$, $g''(t) > 0$, $g'(t) > 0$, and hence $g(t) > 0$. Since $\rho < 1$ implies $t > 0$, we have thus shown that, with $\lambda = 1$ and $\rho < 1$, $\Pi^{SD}(\beta^{SD}) - \Pi^{EPD}(\beta, k) > 0$. (If $\rho = 1$, then $t = 0$, hence $\Pi^{SD}(\beta^{SD}) - \Pi^{EPD}(\beta, k) = 0$.)

Proof of Part (ii): Since Proposition 5 assumes that both EAR and EPD induce interior solutions for efforts, we analyze EAR and EPD separately to prove Part (ii).

We first show that if EPD induces a corner solution for efforts for given (β, k) , then it can be strictly dominated in terms of payoffs by a suitably designed SD scheme. When EPD induces a corner solution for efforts (so $e^{EPD} = 0$), the first-order condition (36) for \bar{e}^{EPD} reduces to (contrast this with (8)):

$$\exp \{r\beta(1-k)\bar{e}^{EPD}\} \frac{\Phi\left(\frac{\bar{e}^{EPD}}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)}{\Phi\left(\frac{-\bar{e}^{EPD}}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)} = \frac{\beta - \bar{e}^{EPD}}{\bar{e}^{EPD} - k\beta}. \quad (48)$$

Since the left-hand side of (48) is strictly greater than 1 for $k < 1$, (48) implies that $\bar{e}^{EPD} < \frac{\beta(1+k)}{2}$. When EPD induces A to choose the corner solution $(\bar{e}^{EPD}, 0)$,

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

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

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

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

$$\begin{aligned} &\frac{r^2\sigma^2}{4}[2\beta^2(1+2\rho k+k^2) - 4(1+\rho)(\bar{e}^{EPD})^2] \\ &\quad + \ln [\Phi(+)+\exp\{-r\beta(1-k)\bar{e}^{EPD}\}\Phi(-)] \\ &\quad + r\beta(1-k)\bar{e}^{EPD}\Phi\left(\frac{-\bar{e}^{EPD}}{\theta}\right) + r\beta\theta(1-k)\phi\left(\frac{\bar{e}^{EPD}}{\theta}\right). \quad (49) \end{aligned}$$

Since (48) implies that $\bar{e}^{EPD} < \frac{\beta(1+k)}{2}$, the expression on the first line of (49) is strictly greater than

$$\frac{r^2\beta^2\sigma^2}{4} [2(1+2\rho k+k^2) - (1+\rho)(1+k)^2]. \quad (50)$$

Using (50), the definitions of y and t in (39), and the second and third lines of (49), we conclude that

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

where $h(y, t)$ is defined in (43). We showed there that $h(y, t)$ is decreasing in y , for all $y \geq 0, t \geq 0$, hence

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

where $g(t)$ was defined and shown in the proof of Part (i) to be strictly positive for all $t > 0$. Therefore, $\Pi^{SD}(\beta^{SD}) - \Pi^{EPD}(\beta, k) > 0$.

An analogous argument shows that if EAR induces a corner solution for efforts for given (β, k) , then it can be strictly dominated in terms of payoffs by a suitably designed SD scheme. This argument starts from the first-order

condition for \bar{e}^{EAR} when the optimal value of $e^{EAR} = 0$ (compare (2)):

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

Proof of Part (iii): The proof of Part (ii) dealt with the case where EAR and EPD induce corner solutions for efforts. From Propositions 5 and 4, we know EPD is more profitable than EAR for any given (β, k) when both schemes induce interior solutions and that EPD induces interior solutions whenever EAR does. Hence, it suffices to show, when $\delta < \lambda$, that for any (β, k) such that EPD induces interior solutions, EPD can be strictly dominated in terms of payoffs by a suitably designed symmetric deterministic (SD) scheme.

From the proof of Proposition 5 (Step 1), we know that we can write

$$\begin{aligned} \Pi^{EPD}(\beta, k) &= \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}h(y, t) \\ &\leq \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}h(0, t) \\ &= \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}[-\ln(2\Phi(t)) + 2t\phi(0)] \\ &< \frac{\bar{e} + \lambda\underline{e}}{\delta} - \frac{1}{2}(\bar{e} + \lambda\underline{e})^2 - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}[-\ln(2\Phi(t)) + 2t\phi(0)] \\ &= \frac{\beta(1+k)}{\delta(1+\lambda)} - \frac{1}{2}\frac{\beta^2}{(1+\lambda)^2} - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}[-\ln(2\Phi(t)) + 2t\phi(0)], \end{aligned} \quad (51)$$

where the weak inequality follows from the fact that $h(y, t)$ is decreasing in y , the strict inequality from the fact that, by assumption, $\delta < \lambda$, and the final equality uses (7).

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 : $\beta^{SD} = \frac{\beta(1+k)}{1+\lambda}$. Since $\lambda > \delta \geq 1$, the SD scheme induces $\bar{e} = \beta^{SD}$, $\underline{e} = 0$, and the principal's payoff under this SD scheme is

$$\Pi^{SD}(\beta^{SD}) = \frac{\beta(1+k)}{\delta(1+\lambda)} - \frac{1}{2}\frac{\beta^2(1+k)^2}{(1+\lambda)^2} - r\sigma^2\beta^2(1+k)^2\frac{(1+\rho)}{(1+\lambda)^2}. \quad (52)$$

Hence from (51) and (52) we can conclude that

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

where the strict inequality is a consequence of the inequalities in (51), the weak inequality follows since $\lambda \geq 1$, the first equality uses (33), the second equality uses (39), and the final line uses the definition of $g(t)$ and its nonnegativity, from the proof of Part (i). ■

B Online Appendix: Not for Publication

B.1 Ex Ante Randomization and the Choice of How Many Tasks to Reward

In Section 7.2 we discussed the trade-offs involved in the design of randomized incentive schemes in environments with many tasks. In this section we provide the derivations for our results.

Consider an EAR scheme in which 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 . Since this scheme is symmetric with respect to all n tasks and since each type of agent's preferences are symmetric with respect to each of his $n - 1$ "non-dislike" tasks, each agent's optimal effort profile can be described by \underline{e} , his effort on his disliked task, and by \bar{e} , his effort on each of the other tasks. If the task that an agent dislikes is included (respectively, not included) in the chosen subset, denote his (conditional) expected utility by \underline{EU} (respectively, \bar{EU}). For any given task, the number of subsets that include it is $\binom{n-1}{\kappa-1}$, while the number that do not is $\binom{n}{\kappa} - \binom{n-1}{\kappa-1} = \binom{n-1}{\kappa}$. Hence each type of agent's unconditional expected utility is

$$\frac{\binom{n-1}{\kappa}}{\binom{n}{\kappa}} \bar{EU} + \frac{\binom{n-1}{\kappa-1}}{\binom{n}{\kappa}} \underline{EU}.$$

We focus on the case where optimal efforts are interior.

The aggregate effort exerted by an agent is $\lambda \underline{e} + (n - 1) \bar{e}$, which we define as A . To find the optimal level of A , we equate the sum over all tasks of the expected marginal monetary returns to effort to the sum over all tasks of the marginal cost of effort. (Formally this corresponds to adding the first-order conditions for effort on each of the n tasks.) This yields $\kappa \beta = (n - 1 + \lambda) A$, so the optimal level of $A = \frac{\kappa \beta}{n - 1 + \lambda}$. To derive the optimal value of $\bar{e} - \underline{e}$, we need the first-order condition for \underline{e} , which is

$$\binom{n-1}{\kappa-1} [\beta - \lambda A] \underline{EU} + \binom{n-1}{\kappa} [-\lambda A] \bar{EU} = 0, \quad (53)$$

since the net marginal monetary return to \underline{e} is $\beta - \lambda A$ if the subset of rewarded tasks includes the agent's disliked one and is $-\lambda A$ otherwise. Substituting for the optimal value of A in (53) and rearranging yields

$$\bar{e} - \underline{e} = \frac{1}{r\beta} \ln \left[\frac{\lambda(n - \kappa)}{n - 1 - (\kappa - 1)\lambda} \right].$$

A necessary condition for interior solutions is $k - 1 \leq \frac{n-1}{\lambda}$. Each type of agent's unconditional expected utility is given by

$$\begin{aligned} EU = & -\frac{\binom{n-1}{\kappa-1}}{\binom{n}{\kappa}} \exp \left\{ -r \left[\alpha + \beta((\kappa - 1)\bar{e} + \underline{e}) - \frac{1}{2} \frac{\kappa^2 \beta^2}{(n - 1 + \lambda)^2} - \frac{1}{2} r \sigma^2 \beta^2 \kappa (1 + \rho(\kappa - 1)) \right] \right\} \\ & - \frac{\binom{n-1}{\kappa}}{\binom{n}{\kappa}} \exp \left\{ -r \left[\alpha + \beta \kappa \bar{e} - \frac{1}{2} \frac{\kappa^2 \beta^2}{(n - 1 + \lambda)^2} - \frac{1}{2} r \sigma^2 \beta^2 \kappa (1 + \rho(\kappa - 1)) \right] \right\}. \end{aligned}$$

The principal will optimally set α to ensure that the participation constraint binds for each type of agent. With α set in this way, and using the expressions for the optimal values of A and $\bar{e} - \underline{e}$, the principal's expected payoff as a function of β and κ can be simplified to

$$\begin{aligned} \Pi(\beta, \kappa) = & \left(\underline{e} + \frac{(n - 1)}{\delta} \bar{e} \right) - \frac{\kappa^2 \beta^2}{2(n - 1 + \lambda)^2} \\ & - \frac{1}{2} r \sigma^2 \beta^2 \kappa (1 + \rho(\kappa - 1)) - \frac{1}{nr} \ln \left[\frac{(n - \kappa)^{n-\kappa} (n - 1 + \lambda)^n}{n^n \lambda^\kappa ((n - 1) - (\kappa - 1)\lambda)^{n-\kappa}} \right], \quad (54) \end{aligned}$$

where

$$\underline{e} + \frac{(n - 1)}{\delta} \bar{e} = \left(\frac{\delta + n - 1}{\delta} \right) \frac{\kappa \beta}{(n - 1 + \lambda)^2} - \frac{(\delta - \lambda)(n - 1)}{\delta(n - 1 + \lambda)r\beta} \ln \left[\frac{\lambda(n - \kappa)}{(n - 1) - (\kappa - 1)\lambda} \right]. \quad (55)$$

Using $\tilde{\beta} = \kappa \beta$ to substitute for β in the above payoff expression yields expressions (18) and (19) in the text.

Verifying the claims in the final paragraph of Section 7.2 regarding the effect of varying δ , r (with $r\sigma^2$ fixed), or $\sigma^2(1 - \rho)$, on the optimal value of κ requires signing the cross-partial derivative of $\Pi(\tilde{\beta}, \kappa)$ in (18) with respect to κ and the relevant parameter, holding $\tilde{\beta}$ fixed. We can show that $\frac{\partial^2 \Pi}{\partial \delta \partial \kappa} < 0$, $\frac{\partial^2 \Pi}{\partial r \partial \kappa} > 0$, and $\frac{\partial^2 \Pi}{\partial (\sigma^2(1-\rho)) \partial \kappa} > 0$, from which the claims follow. The final claim follows from the fact that $\frac{\partial^2 \Pi}{\partial \tilde{\beta} \partial \kappa} > 0$.