CIFR
Centre for International Finance and Regulation
Towards Financial System Integrity

Research Working Paper Series

Gaming and Strategic Opacity in Incentive Provision

Professor Florian Ederer
Assistant Professor of Economics
Yale School of Management

Professor Richard Holden
Australian School of Business
UNSW Australia

Dr Margaret Meyer
Official Fellow in Economics
Nuffield College
University of Oxford

WORKING PAPER NO. 007/2014
MARCH 2014

www.cifr.edu.au

This research was supported by the Centre for International Finance and Regulation (project number E225) which is funded by the Commonwealth and NSW Governments and supported by other Consortium members (see www.cifr.edu.au).

GAMING AND STRATEGIC OPACITY IN INCENTIVE PROVISION

By

Florian Ederer, Richard Holden, and Margaret Meyer

January 2014

COWLES FOUNDATION DISCUSSION PAPER NO. 1935

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Gaming and Strategic Opacity in Incentive Provision*

Florian Ederer† Richard Holden‡ Margaret Meyer§
Yale UNSW and NBER Oxford and CEPR

January 20, 2014

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 in a two-task moral hazard model in which the agent is privately informed about which task is less costly for him to work on. We examine two simple classes of incentive scheme that are "opaque" in that they make the agent uncertain ex ante about the values of the incentive coefficients in the linear payment rule. We show that, relative to deterministic menus of linear contracts, these opaque schemes induce more balanced efforts, but they also impose more risk on the agent per unit of aggregate effort induced. We identify settings in which optimally designed opaque schemes not only strictly dominate the best deterministic menu but also completely eliminate the efficiency losses from the agent's better knowledge of the environment. Opaque schemes are more likely to be preferred to transparent ones when i) efforts on the tasks are highly complementary for the principal; ii) the agent's privately known preference between the tasks is weak; iii) the agent's risk aversion is significant; and iv) the errors in measuring performance on the tasks have large correlation or small variance. (*JEL* D86, D21, 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,

*We thank Philippe Aghion, Sushil Bikhchandani, Simon Board, Patrick Bolton, Jeremy Bulow, Vince Crawford, Kate Doornik, Guido Friebel, Robert Gibbons, Edward Glaeser, Oliver Hart, Bengt Holmström, Ian Jewitt, Navin Kartik, Paul Klemperer, Shmuel Leshem, Phillip Leslie, Steven Lippman, Steven Matthews, John Moore, Andy Newman, In-Uck Park, Charles Roddie, Jesse Shapiro, Ferdinand von Siemens, Jason Snyder, Alexander Stremitzer, Jeroen Swinkels, and Jean Tirole for helpful discussions and suggestions, as well as seminar participants at Bonn, Boston University, Bristol, Chicago, Columbia, Frankfurt, Gerzensee, LSE, MIT, Oxford, Rotterdam, UCL, UCLA, USC, and Yale. Holden thanks the Centre for International Financial Regulation at UNSW for financial support.

†Yale School of Management, Yale School of Management, 165 Whitney Avenue, New Haven, CT 06511, U.S.A., florian.ederer@yale.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.

performance; and iii) distortion of choices about timing 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), Larkin (2013), 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–deliberate opacity about the criteria upon which rewards will be based and/or how heavily these criteria will be weighted–can help deter gaming. This idea has a long intellectual history. It dates back at least to Bentham (1830), who argued that deliberate opacity about the content of civil service 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.

The costs of transparency have also been discussed in the context of gaming, by law school deans, of the performance indicators used by *U.S. News* to produce its highly influential law school rankings. The ranking methodology is transparent and employs a *linear* scoring rule.[3] Law scholars (e.g. Osler, 2010) have argued that greater opacity in the ranking methodology could mitigate gaming, and *U.S. News* has itself signaled its intention to move away from being "totally transparent about key methodology details".[4] [5]

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

[1]Burgess 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) use data from postsecondary education, while Oyer (1998), Larkin (2013) and Forbes et al (2012) examine private sector organizations (salespeople and executives across various industries, enterprise software vendors, and airlines, respectively).

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

[3]The weights in the scoring rule are quality perception (40%), selectivity (25%), placement success (20%) and faculty resources (15%) (U.S. News, March 11, 2013, `http://www.usnews.com/education/best-graduate-schools/top-law-schools/articles/2013/03/11/methodology-best-law-schools-rankings`).

[4]U.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`.

[5]Relatedly, Google has experienced manipulation of its search results by some retailers. Although many retailers have been seeking greater transparency from Google about its search algorithm, Google has responded by moving in the direction of greater opacity to prevent manipulation (Structural Search Engine Optimization, Google Penalty Solutions, November 4, 2011, `http://www.re1y.com/blog/occupy-google-blog.html`). Jehiel and Newman (2011) develop a dynamic model in which principals learn from agents' behavior about the possibilities for gaming of incentive schemes and then choose whether to take costly measures to deter gaming.

can improve parties' incentives to fulfill the spirit of the contract rather than focusing on satisfying only 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, in order to encourage somewhat slower driving everywhere, 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 low, subjects are often paid based on their performance in one randomly chosen period of the experimental session.[6]

The settings discussed above all suggest that opacity of incentive schemes can be beneficial in reducing gaming, especially where incentive designers care about multiple aspects of agents' performance and gaming takes the form of focusing efforts on easily manipulable indicators. This line of argument is, however, incomplete. If agents are risk averse, then the additional risk imposed by opaque schemes is per se unattractive to them. Understanding when and why opaque schemes are used requires analyzing the tradeoff between their incentive benefits and their risk costs. This paper provides such an analysis.

We develop a formal model of gaming by an agent with superior knowledge of the environment, and we explore when and to what extent "opacity", i.e. lack of transparency about the weighting scheme used to determine rewards, can mitigate gaming. A 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.[7] 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. The agent games transparent reward schemes by choosing effort allocations that are excessively (from an efficiency perspective) sensitive to his private information. In fact, we show that the agent's superior knowledge of his preferences makes it impossible for the principal, with transparent schemes, to induce both types of agent to exert positive efforts on both tasks, even when menus of contracts are used as screening devices.

In this setting, we study the performance of two simple types of "opaque" incentive scheme. Each scheme is opaque in that, while the agent knows that the compensation schedule ultimately used will take one of two possible linear forms, *at the time he chooses his efforts he does not know which form will be used.* The two possible compensation schedules differ with respect to which performance measure is more highly rewarded. Under 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. Although ex ante randomization and ex post discretion differ with respect to the mechanism determining which of the two possible linear schedules will ultimately be used, they are both opaque in that they make the agent uncertain

[6]Take 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)

[7]The analysis would be very similar if the agent types differed with respect to the task on which they were more productive.

ex ante about the incentive coefficients in the linear payment rule. We show that this ex ante uncertainty of the agent produces qualitatively similar incentive effects under the two types of opaque scheme.[8]

Ex ante randomization encourages the risk-averse agent to choose relatively 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 opaque incentive scheme, we show that the more unequal are the weights on the performance measures in the two possible compensation schedules, the stronger are the agent's incentives to choose balanced efforts.

We demonstrate that the performance of opaque incentive schemes is more robust to uncertainty about the agent's preferences than is the performance of deterministic ones. With opaque schemes, the efforts exerted on the two tasks vary continuously when a small amount of such uncertainty is introduced, whereas they vary discontinuously for deterministic ones.

The benefits of opaque incentive schemes in deterring gaming do, nevertheless, come at a cost, as mentioned above: such schemes impose more risk on the agent. While ex post discretion can be shown to impose lower risk than ex ante randomization, still, for any level of *aggregate* effort induced on the two tasks, we show that a deterministic contract imposes lower risk costs than either type of opaque scheme. As a consequence, as we prove in Proposition 6, if the principal faces no uncertainty about the agent's preferences, then both types of opaque 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 opaque schemes and the lower risk costs under deterministic ones.

Our key contribution is to identify settings in which both of our simple types of opaque incentive scheme, when designed optimally, strictly dominate all deterministic linear menus of contracts. We identify three such environments. In each one, optimally weighted opaque 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 becomes very small. The second is the case where the agent's risk aversion becomes very large and the variance of the shocks to outputs becomes very small. In the final setting, the correlation of the output shocks becomes very high. 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 the agent's hidden information. That is, in these settings, both simple types of opaque incentive scheme completely eliminate the efficiency losses from the agent's better knowledge of the environment.

Though our propositions focus on limiting environments to prove analytically that ex ante randomization and ex post discretion can strictly dominate all deterministic linear menus, our results have more general implications about what characteristics of contracting environments increase the relative attractiveness of opaque schemes. Opaque schemes are more likely to be preferred when i) efforts on

[8] The term "opaque" may have alternative definitions in other contexts, but in this paper, it will be used exclusively in the sense just defined.

the tasks are highly complementary for the principal; ii) the agent's privately known preference between tasks is weak, so even a small amount of uncertainty about the weights in the compensation schedule has a large effect on how balanced the agent's chosen efforts are; iii) the agent's risk aversion is significant, so opaque schemes provide the agent with a powerful self-insurance motive for balancing efforts; and iv) the errors in measuring performance on the tasks have large correlation or small variance.

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).[9] Motivated not only by these robustness arguments, but also by the simplicity and pervasiveness of linear contracts, 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 Malcomson (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 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; the attractiveness of ex ante randomization and ex post discretion stems from their ability to mitigate the agency costs of *multi-task* incentive problems when compensation schedules are constrained to be ex post linear.

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,

[9] Carroll (2012) also demonstrates an appealing robustness property of linear contracts. He shows that, in a static model with limited liability, when the principal knows some but not all of the actions available to the agent and evaluates contracts according to their worst-case performance, a linear contract is optimal.

Possen, 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 7.5.

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 Lazear's analysis of testing, there is an exogenous restriction on the number of topics that can be tested, whereas in our model, even when all tasks can be measured and rewarded, we show that deliberate opacity about the weights in the incentive scheme can be desirable.

The costs and benefits of transparency are also a focus of interest in international relations. Baliga and Sjöström (2008) show in a model of arms proliferation that a small country may be able to use what they term "strategic ambiguity" about whether it possesses weapons of mass destruction as a substitute for actual investment and thereby help to deter an attack. Wikipedia defines the policy of "strategic ambiguity" 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."[10] Multiple objectives of the principal and risk aversion of the agent are also important in our model in generating the beneficial incentive effects of opacity.[11]

The model perhaps most closely related to ours is that of 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 in their model 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, under the specific assumptions of their model, ex ante randomization over separable payment schemes would have no power to mitigate gaming, because even for risk averse agents it would not generate a self-insurance motive for

[10] Wikipedia, "Policy of Deliberate Ambiguity", `http://en.wikipedia.org/wiki/Policy_of_deliberate_ambiguity`.

[11] It is important to stress that our results on the benefits of opacity do not rely on Knightian uncertainty or on the presence of ambiguity-averse agents.

choosing more balanced efforts. In contrast, our analysis reveals (see Section 7.3) that even beyond the exponential-normal setting on which we focus, ex ante randomization will generate a self-insurance motive and will thereby mitigate the excessive sensitivity of agents' effort allocations to their private information.

Section 2 outlines the model, and Section 3 studies deterministic incentive schemes. Section 4 analyzes our two classes of opaque schemes, ex ante randomization and ex post discretion. Section 5 identifies settings in which opaque schemes are dominated by deterministic ones. Section 6, which is the heart of the paper, identifies environments in which optimally weighted opaque schemes dominate the best deterministic one. Section 7 discusses the robustness of our results as well as some extensions, 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).[12] Nevertheless, since $\lambda \geq 1$, the type-i agent has a preference for task i: the marginal cost of effort on task j ($j \neq i$) is λ times as large as that on task i. 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 substitutes—here the principal is indifferent as to how the agent allocates his total effort across the tasks.[13]

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

[12] In Section 7.2, we show that our key results hold even when the degree of substitutability of efforts for the agent is high but imperfect.

[13] We 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).

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 *opaque*.

In the next section, we study deterministic incentive schemes. Section 4 then analyzes the performance of the two simple classes of opaque 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 according to $w = \alpha + \beta x_1 + k\beta x_2$ or $w = \alpha + \beta x_2 + k\beta x_1$, where again $k \in (-1, 1)$. Under both classes of opaque incentive schemes, the agent is ex ante uncertain about what weights the two performance indicators will be given in the linear formula determining his pay, but under ex post discretion the agent's efforts themselves influence which set of weights is ultimately used.

3 Deterministic Contracts

3.1 The No Hidden Information Benchmark

Suppose that 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). The NHI benchmark is important because, as we will see, there are environments in which optimally designed opaque 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. The first form is a pair of contracts (C_1^{bal}, C_2^{bal}), where

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

and where the principal assigns the contract C_i^{bal} to the type-i agent. The incentive coefficients in C_i^{bal} are chosen to equate the ratio of the marginal benefits of efforts on the two tasks to the ratio of their marginal costs for type i. As stressed by Holmstrom and Milgrom (1991) and Milgrom and Roberts (1992, p.228), equalizing these ratios is necessary for a contract to induce strictly positive efforts on both tasks, an observation often referred to as the "equal compensation principle". Here, since these ratios are constant, independent of the chosen efforts, it follows that type i is indifferent over all non-negative effort pairs satisfying $\beta = e_i + \lambda e_j$. Among such effort pairs, the principal prefers type i to choose $e_i = e_j = \frac{\beta}{1+\lambda}$, since efforts on the tasks are complementary for the principal, and we assume that type i does indeed choose this perfectly balanced effort allocation.

In the special case where $\lambda = 1$, there is only one type of agent, and C_1^{bal} and C_2^{bal} both reduce to the "symmetric deterministic" (SD) contract

$$SD : w = \alpha + \beta x_1 + \beta x_2.$$

When $\lambda = 1$, the SD contract makes the agent willing to choose $e_1 = e_2 = \frac{\beta}{2}$. In the no hidden information (NHI) benchmark, the efforts induced by the contract pair (C_1^{bal}, C_2^{bal}), and hence the payoff received by the principal, are continuous in λ, approaching their values under the SD contract as $\lambda \to 1$.

The second type of contract pair which can be optimal in the NHI benchmark is a pair of the form

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

where the principal assigns C_i^{foc} to the type-i agent. Contract C_i^{foc} induces type i to exert effort only on his preferred task, task i, and to set $e_i = \beta$ and $e_j = 0$, for any $\lambda \geq 1$. Contract C_i^{foc} uses performance on task j to provide insurance for the type-i agent (without weakening his incentives on task i), by exploiting the correlation between the shocks to the two performance measures. Among all contract pairs that induce each type to focus only on his preferred task, pairs of the form (C_1^{foc}, C_2^{foc}) are the most attractive for the principal.[14]

In choosing, in the NHI setting, between a contract pair of the form (C_1^{bal}, C_2^{bal}) and one of the form (C_1^{foc}, C_2^{foc}), the principal faces a trade-off between the more balanced efforts induced by the former and the lower risk cost imposed by the latter. If and only if the efforts on the two tasks are sufficiently complementary for the principal, the benefits of the balanced efforts elicited by (C_1^{bal}, C_2^{bal}) outweigh the costs of the extra risk imposed on the agent by this contract pair.

Lemma 1 *For any $\lambda \geq 1$, in the no hidden information (NHI) benchmark, there exists a critical value of the task complementarity parameter δ in the principal's benefit function, $\delta^{NHI}(\lambda, r\sigma^2, \rho)$, increasing in each of its arguments, such that for $\delta > \delta^{NHI}$ (respectively, $\delta < \delta^{NHI}$), the principal's unique optimal contract pair has the form (C_1^{bal}, C_2^{bal}) (respectively, the form (C_1^{foc}, C_2^{foc})).*

[14] Although the values of α and β could in principle be allowed to differ between C_1^{bal} and C_2^{bal} and, analogously, between C_1^{foc} and C_2^{foc}, the symmetry of the model with respect to the two types of agent makes it optimal for these values to be the same within each type of contract pair. Moreover, this symmetry also implies that it is never uniquely optimal to offer a pair of the form (C_1^{foc}, C_2^{bal}) or (C_1^{bal}, C_2^{foc}).

3.2 The General Case: Hidden Information

When $\lambda > 1$ and the agent is privately informed about his preferences across tasks, the principal can use menus of contracts as a screening device.

Proposition 1 *(i) When $\lambda > 1$ and the agent's type is hidden information, no menu of linear deterministic contracts can induce both types of agent to choose strictly positive efforts on both tasks. (ii) For any $\lambda > 1$, if $\delta > \delta^{NHI}(\lambda, r\sigma^2, \rho)$, the principal is strictly worse off when hidden information is present than when it is absent. (iii) For $\delta > \delta^{NHI}(1, r\sigma^2, \rho)$, the limit as λ approaches 1 of the principal's maximized payoff under hidden information is strictly less than in the no hidden information benchmark.*

To prove part (i), we begin by observing that the "equal compensation principle" has the following implication for a menu of deterministic linear contracts: the only way to induce both types of agent to exert strictly positive efforts on both tasks is to induce each type to choose a contract that equates the ratio of the marginal benefits of efforts on the tasks to the ratio of their marginal costs. Therefore, if a menu existed which could induce both types to choose strictly positive efforts on both tasks, it would have the form

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

and would induce the type-i agent to choose contract C_i. 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.

Part (iii) of the proposition strengthens the result in part (ii) to show that the principal's maximized payoff under hidden information is not only strictly below that in the NHI benchmark but is also bounded away from it, even as λ approaches 1. Parts (ii) and (iii) follow from part (i), combined with the result from Lemma 1 that, in the NHI benchmark, for any $\lambda \geq 1$ and for $\delta > \delta^{NHI}(\lambda, r\sigma^2, \rho)$, the principal *strictly* prefers to induce both types of agent to choose strictly positive (in fact, perfectly balanced) efforts on the two tasks.[15]

4 Opaque Contracts

4.1 Ex Ante Randomization

A contract with *ex ante randomization* (EAR) specifies that with probability $\frac{1}{2}$, the agent will be compensated according to $w = \alpha + \beta x_1 + k\beta x_2$, and with probability $\frac{1}{2}$, he will be compensated according to $w = \alpha + \beta x_2 + k\beta x_1$, where the key parameters are the incentive intensity $\beta > 0$ and the weighting factor $k \in (-1, 1)$. (The lump-sum payment α has no effect on the agent's incentives, and will always be set by the principal to make the participation constraint binding for both types of agent.) Under this incentive 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

[15] It can be shown that there is a critical value of δ above which the optimal deterministic menu induces one type of agent to choose perfectly balanced efforts and the other type to choose fully focused efforts, and below which the optimal menu induces both types to choose fully focused efforts. The details of these contracts are, however, not needed for the analysis that follows.

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 defined in Section 3.1, 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, $\overline{e}^{EAR}$, and effort on his more costly task, $\underline{e}^{EAR}$, satisfying

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

$$\exp\left[r\beta(1-k)(\overline{e}^{EAR} - \underline{e}^{EAR})\right] = \frac{\lambda - k}{1 - k\lambda}; \tag{2}$$

(iii) the gap in efforts, $\overline{e}^{EAR} - \underline{e}^{EAR}$, is increasing in λ, approaching 0 as $\lambda \to 1$; decreasing in $r\beta$, approaching 0 as $r\beta \to \infty$; and increasing in k, approaching 0 as $k \to -1^{+}$;
(iv) the principal's payoff under EAR, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$, is

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

To understand part (i), note that if $k \geq \frac{1}{\lambda}$, then for both types of agent, *whichever of the two compensation schedules is randomly selected*, the ratio of the marginal benefit of effort on the preferred task to that on the less-preferred task is at least as large as the corresponding ratio of the marginal costs of effort, and strictly larger for one of the schedules. It follows from the "equal compensation principle", therefore, that both types of agent would optimally exert effort only on their preferred task.

To understand equation (1), note first that the sum of the marginal monetary returns to effort on the two tasks is certain to be $\beta(1+k)$, since regardless of the outcome of the randomization, 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 $\overline{e}$ and $\underline{e}$ must yield $\beta(1+k) = \partial c/\partial \overline{e} + \partial c/\partial \underline{e}$ for both types of agent. Since $\partial c/\partial \overline{e} + \partial c/\partial \underline{e} = (1+\lambda)(\overline{e} + \lambda \underline{e})$, this gives us (1). Throughout, we will refer to the quantity $\overline{e} + \lambda \underline{e}$ as an agent's *aggregate effort*, since it is the quantity which determines his total cost of effort.

To derive equation (2), observe that, by the "equal compensation principle", if optimal efforts are interior, then the ratio of the expected marginal-utility-weighted marginal benefits of the two types of effort must equal the corresponding ratio of their marginal costs. Equating these ratios, and using the fact that each of the two possible compensation schedules is employed with probability one-half, yields

$$\frac{kE\left[U'(\cdot)I_{\{\overline{x}\text{ is rewarded more highly than }\underline{x}\}}\right] + E\left[U'(\cdot)I_{\{\underline{x}\text{ is rewarded more highly than }\overline{x}\}}\right]}{E\left[U'(\cdot)I_{\{\overline{x}\text{ is rewarded more highly than }\underline{x}\}}\right] + kE\left[U'(\cdot)I_{\{\underline{x}\text{ is rewarded more highly than }\overline{x}\}}\right]} = \lambda, \tag{4}$$

which simplifies to

$$\frac{E\left[U'(\cdot)I_{\{\underline{x}\text{ is rewarded more highly than }\overline{x}\}}\right]}{E\left[U'(\cdot)I_{\{\overline{x}\text{ is rewarded more highly than }\underline{x}\}}\right]}=\frac{\lambda-k}{1-k\lambda}, \tag{5}$$

where $\overline{x}$ (respectively, $\underline{x}$) denotes performance on an agent's less costly (respectively, more costly) task. The left-hand-side of (5) reduces to

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

thus yielding (2).

The equations above reveal why ex ante randomization can provide both types of agent with incentives to choose positive efforts on both tasks. Because of the uncertainty about which compensation schedule will be used, and because of the agent's risk aversion, the left-hand-side of (4), which equals the ratio of the expected marginal utility gain from increasing $\underline{e}$ to that from increasing $\overline{e}$, is not a constant; rather, it is a continuously increasing function of the gap between efforts on the two tasks, $\overline{e}-\underline{e}$.[16] In other words, as the effort allocation contemplated by the agent becomes more focused on his less costly task (i.e., as $\overline{e}-\underline{e}$ increases), the greater is the risk he faces from the random choice of compensation schedule, and the relatively more attractive it becomes to self-insure by raising effort on his more costly task.

Equation (2) shows how the agent's optimal degree of self-insurance against the compensation risk imposed by EAR varies with his preferences and with the parameters of the incentive scheme. 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 effort gap $\overline{e}-\underline{e}$. The more risk-averse the agent (the larger is r) or the larger the incentive intensity β, the more costly 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 $\lambda\to 1$ or $r\beta\to\infty$, 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 $k\to -1^+$, the self-insurance motive approaches its strongest level, and the optimal effort gap approaches 0.

Under the symmetric deterministic contract defined in Section 3.1, 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, Proposition 1 shows that, even when the principal chooses a deterministic linear menu optimally as a function of λ, his payoff drops discontinuously as λ is increased from 1 (as long as $\delta>\delta^{NHI}(1,r\sigma^2,\rho)$, so that inducing balanced efforts from both types of agent would, if feasible, actually be strictly optimal). 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

[16]In contrast, if the agent were certain which compensation schedule would be used, the left-hand-side of (4) would reduce to either k or $\frac{1}{k}$.

best deterministic menu.[17] EAR is also more robust to uncertainty about the magnitude of λ than is a deterministic menu. If the principal tries to design a 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.[18]

We have established Proposition 2 under the assumption that the principal can commit to randomizing uniformly between the two compensation schedules.[19] 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.[20] Thus the attractive properties of ex ante randomization are not crucially dependent on the principal's having the power to commit to the randomizing probability.

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 $\overline{e} + \lambda\underline{e}$ 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:

$$\Pi^{EAR}(B, k) = \underline{e}^{EAR} + \frac{\overline{e}^{EAR}}{\delta} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2}r\sigma^2 B^2 \frac{1 + 2\rho k + k^2}{(1+k)^2} - \frac{1}{2r}\ln\left[\frac{(\lambda+1)^2(1-k)^2}{4(1-k\lambda)(\lambda-k)}\right], \quad (6)$$

where

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

[17] In Section 7.3 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.

[18] Bond 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".

[19] Given the power to commit to a randomizing probability, it is optimal for the principal to commit to randomize uniformly. Doing so results in the most balanced profile of effort choices, assessed ex ante, and also avoids leaving any rent to either type of agent.

[20] To see that the outcome described in Proposition 2 is *an* equilibrium under interim randomization, note that given that the two types of agent are equally likely and given that their effort 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 optimal behavior for each type of agent is clearly as described in the proposition. To see that this outcome is the *unique* equilibrium, observe that if the two types of agent conjectured that the principal would choose the schedule rewarding task 1 more highly than task 2 with a probability greater than (less than) 1/2, then their optimal efforts would be such that the principal would strictly prefer to choose the schedule rewarding task 2 more (less) highly than task 1.

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 (6) and (7) show that increasing k has three effects. First, a larger k raises the effort gap $\overline{e} - \underline{e}$ and, with B and hence aggregate effort $\overline{e} + \lambda\underline{e}$ held fixed, this larger gap lowers the principal's benefit $\underline{e} + \frac{\overline{e}}{\delta}$ whenever $\delta > \lambda$, i.e., whenever 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 (6). 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 (6).

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);
(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);
(iv) the smaller is B (i.e., the smaller the level of aggregate effort to be induced).

In Section 6, where we identify environments where optimally weighted EAR outperforms the best deterministic menu, we will build on these results. In Section 7.4, where we study EAR in a setting with an arbitrary number n of tasks, we show that changes in the number of randomly chosen tasks to reward have the same qualitative effects on incentives and risk as do 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 opaque contracts.

4.2 Ex Post Discretion

Under a contract involving *ex post discretion* (EPD), the principal, *after* observing x_1 and x_2, chooses whether to pay the agent according to $w = \alpha + \beta x_1 + k\beta x_2$ or $w = \alpha + \beta x_2 + k\beta x_1$, where as with EAR, the key parameters are $\beta > 0$ and $k \in (-1, 1)$. Just as under EAR, the agent is uncertain at the time he chooses his efforts whether his pay will be more sensitive to performance on task 1 or task 2, but with 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, $\overline{e}^{EPD}$, and effort on his more costly task, $\underline{e}^{EPD}$, satisfying

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

$$\exp\left[r\beta(1-k)d\right] \frac{\Phi\left(\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)}{\Phi\left(\frac{-d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)} = \frac{\lambda - k}{1 - k\lambda}, \tag{9}$$

where $\theta \equiv \sigma[2(1-\rho)]^{\frac{1}{2}}$ and $d \equiv \overline{e}^{EPD} - \underline{e}^{EPD}$;
(iii) the gap in efforts, $\overline{e}^{EPD} - \underline{e}^{EPD}$, is increasing in λ, approaching 0 as $\lambda \to 1$; decreasing in $r\beta$, approaching 0 as $r\beta \to \infty$; increasing in $\sigma^2(1-\rho)$, approaching 0 as $\sigma^2(1-\rho) \to 0$; and increasing in k, approaching 0 as $k \to -1^+$;
(iv) the principal's payoff under EPD, for given $\beta > 0$ and $k \in (-1, \frac{1}{\lambda})$, is

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

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 $\overline{e}^{EPD} - \underline{e}^{EPD} < \overline{e}^{EAR} - \underline{e}^{EAR}$.

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, $\overline{e} + \lambda\underline{e}$, is the same under EPD as under EAR—compare equations (8) and (1).[21] Since both schemes are certain to reward one task at rate β and the other at rate $k\beta$, the sum of the expected marginal returns to effort on the two tasks is $(1+k)\beta$ under both schemes, and for interior solutions, this sum is equated to the sum of the marginal effort costs on the two tasks, $(\lambda+1)(\overline{e} + \lambda\underline{e})$. Just as for EAR, it follows from the "equal compensation principle" that if optimal efforts are interior, then equation (4), and hence equation (5), must hold. However, for EPD

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

[21]Intuitively, 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, $\overline{e} + \underline{e}$, is lower under EPD than under EAR. Nevertheless, aggregate effort $\overline{e} + \lambda\underline{e}$, and hence the costs of effort incurred, are the same under the two schemes.

which when combined with (5) gives us (9).

Under EAR, the risk-averse agent's incentive to choose (partially) balanced efforts derives purely from an insurance motive: a desire to insure himself against the risk generated by the random choice of compensation schedule. Under EPD, 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 (9), which is increasing in $(\overline{e}-\underline{e})$, is strictly greater than the left-hand side of (2) for all $(\overline{e}-\underline{e}) > 0$, and this implies that for all $\lambda > 1$, $\overline{e}^{EPD}-\underline{e}^{EPD} < \overline{e}^{EAR}-\underline{e}^{EAR}$.

Under EPD, the agent's optimal choice of $\overline{e} - \underline{e}$ 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 $\lambda \to 1$ or $r\beta \to \infty$, $(\overline{e} - \underline{e}) \to 0$. 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 $\overline{e} - \underline{e}$ more likely to affect which compensation schedule the principal chooses, so gives the agent a stronger incentive to balance his efforts. As $\sigma^2(1-\rho) \to 0$, for example because the shocks become perfectly correlated, optimal efforts become perfectly balanced.

Under EPD, just as under EAR, reducing the parameter k, and so making the two possible compensation schedules more different, induces the agent to choose more balanced efforts. While the effect of k on the effort gap $(\overline{e} - \underline{e})$ is more complex under EPD than under EAR, nevertheless the 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 (10), 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 a *lower* total 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 a risk cost $\frac{1}{2}r\sigma^2\beta^2(1+2\rho k+k^2)$.[22] The intuitive reason for this finding 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 a *higher* total 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 (10) 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 $\underline{e} + \frac{\overline{e}}{\delta}$. On the other hand, a larger k improves the diversification of the measurement errors. Finally, as k and hence $\overline{e} - \underline{e}$ 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 quantitative 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

[22] This is proved formally in Step 1 of the proof of Proposition 5.

induced, will be smaller, the more complementary are the tasks for the principal (i.e., the larger is 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 $\overline{e} - \underline{e}$ than EAR, while the two schemes induce the same aggregate effort $\overline{e} + \lambda\underline{e}$ and hence the same total cost of effort, and (ii) EPD imposes lower risk costs on the agent than EAR. 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 opaque incentive scheme are strictly dominated by a deterministic contract.

Proposition 6 *For given (β, k), where $k \in (-1, 1)$, both EAR and EPD yield a strictly lower payoff for the principal than a suitably designed symmetric deterministic (SD) contract, if any of the following conditions hold:*
(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 each part of this proposition is the important result that, for any weighting factor $k < 1$, and for any λ, the total risk cost imposed by EAR and EPD in inducing the agent to exert any given level of *aggregate effort* $\overline{e} + \lambda\underline{e}$ is larger than that imposed by a SD contract. At the same time, whenever $\lambda > 1$, neither a SD contract nor any deterministic menu can induce the agent to exert positive effort on both tasks (as shown by Proposition 1), whereas EAR and EPD have the potential to induce better-balanced efforts. In general, therefore, the principal faces a trade-off in choosing between the opaque schemes (EAR and EPD) and deterministic ones. Opaque schemes are typically better at inducing balanced efforts, while deterministic ones have the advantage of imposing a lower risk cost on the agent per unit of aggregate effort induced.

The three conditions identified in Proposition 6 are ones under which this trade-off does not in fact arise. Under condition (ii), even the opaque schemes induce corner solutions for efforts. Corner solutions arise, for example, when λ, measuring the agent's bias towards his preferred task, is sufficiently large. Corner solutions also arise when, holding $r\sigma^2$ fixed, r gets sufficiently small (so the self-insurance motive for balancing efforts 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.

The most significant of the three results in Proposition 6 is the first one. If $\lambda = 1$, then the principal faces no uncertainty about the agent's preferences. In this case, the SD contract, EAR, and EPD all induce perfectly balanced efforts. The proof of part (i) shows that, when $\lambda = 1$, 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, so is minimized at $k = 1$, when β is adjusted to keep efforts unchanged.[23] Since the SD contract corresponds to $k = 1$, the SD contract therefore strictly dominates EPD. The implication of part (i) of Proposition 6 is that the agent's superior knowledge of the environment (here, of his preferences, as reflected by a value of $\lambda > 1$) is necessary for EAR and EPD to have the potential to dominate deterministic menus.

6 When Are Opaque Incentive Schemes Optimal?

We now identify three environments in which opaque schemes, when designed optimally, strictly dominate the best linear deterministic menu. In each of these environments, both EAR and EPD, with the weighting 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 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 opaque schemes eliminate the efficiency losses from the agent's hidden information.

6.1 Very Weak Preferences across Tasks for the Agent

Consider first a setting in which the agent has private information about his preferences, but the magnitude of his preference across tasks is very weak. Formally, we study the case in which λ is strictly greater than but arbitrarily close to 1, which we term the limiting case as $\lambda \to 1^+$.

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 $\lambda = 1$. This robustness of EAR and EPD to the introduction of hidden information underlies the superiority of these schemes in the limiting case as $\lambda \to 1^+$, as we now show.

Propositions 2 and 4 show that as $\lambda \to 1$, so the two tasks become equally costly, $(\overline{e} - \underline{e}) \to 0$ for any $k \in (-1, 1)$, under both EAR and EPD. Equations (6) and (7) 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 $\lambda \to 1$, the situation is dramatically simpler:

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

Because, as $\lambda \to 1$, efforts under EAR become perfectly balanced, the risk cost imposed by the

[23] The 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 $(x_2 + kx_1)/(1 + k)$.

randomization 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 (11). Thus, as $\lambda \to 1$, $\Pi^{EAR}(B,k)$ is increasing in k (strictly so for $\rho < 1$), as long as k induces interior solutions, which it does as long as $k < \frac{1}{\lambda}$. Therefore, as $\lambda \to 1$, $\Pi^{EAR}(B,k)$ is maximized, for any B, by setting k arbitrarily close to, but less than, 1 ($k \to 1^-$). With k set in this way, the principal's payoff approaches

$$\lim_{k\to 1}\lim_{\lambda\to 1}\Pi^{EAR}(B,k) = \frac{(\delta+1)}{\delta}\frac{B}{4} - \frac{B^2}{8} - \frac{1}{4}r\sigma^2 B^2 (1+\rho). \qquad (12)$$

The right-hand side of (12) 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 $k \to 1^-$.[24] Thus, for any B, as $\lambda \to 1^+$, 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 states that EPD yields the principal a payoff at least as large as does EAR whenever both schemes induce interior effort choices. Part (i) of Proposition 6 shows that when $\lambda = 1$, the principal's payoff under EPD, for any $k < 1$ and for any B, is less than his payoff under the SD contract inducing the same levels of effort. Such a SD contract has $\beta = \frac{B}{2}$ and generates a payoff given by the right-hand side of (12). It therefore follows that as $\lambda \to 1^+$, optimally designed EPD, like EAR, sets k arbitrarily close to, but less than, 1 ($k \to 1^-$), and for any B yields a payoff arbitrarily close to that from a SD contract at $\lambda = 1$.

For the no hidden information benchmark, Section 3.1 shows that 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. Lemma 1 shows that as $\lambda \to 1$, a pair of the form (C_1^{bal}, C_2^{bal}) is strictly optimal for the principal as long as $\delta > \lim_{\lambda\to 1}\delta^{NHI}(\lambda, r\sigma^2, \rho)$. On the other hand, Proposition 1 shows that under hidden information, even as $\lambda \to 1^+$, the principal's maximized payoff from a deterministic menu is bounded away from that in the NHI benchmark, because even for λ arbitrarily close to 1, it is impossible to induce positive efforts on both tasks from both types of agent.

The arguments in the preceding paragraphs together imply:

Proposition 7 *Consider the limiting case as $\lambda \to 1^+$. Under both EAR and EPD, for any given level of aggregate effort, $\overline{e} + \lambda \underline{e}$, to be induced:*
(i) the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any $k \in (-1, 1)$;
(ii) the optimal value of $k \to 1^-$;
(iii) with k adjusted optimally, the principal's payoff under both EAR and EPD 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 $\delta > \lim_{\lambda\to 1}\delta^{NHI}(\lambda, r\sigma^2, \rho)$, EAR and EPD with k and β adjusted optimally strictly dominate the best deterministic menu under hidden information.

The preceding analysis also has another important implication. Even if $\lambda = 1$, the symmetric deterministic contract leaves the agent indifferent over how any total effort is split between the tasks,

[24] See equation (24) in the proof of Lemma 1 in the appendix, and set $\lambda = 1$.

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 Very Large Risk Aversion and Very Small Variance of the Shocks

Consider now the case where the agent's coefficient of absolute risk aversion, r, gets very large and the variance of the shocks to measured performance, σ^2, gets very small, holding $r\sigma^2$ fixed at $R < \infty$. Propositions 2 and 4 show that, in this environment, for any $k \in (-1, \frac{1}{\lambda})$, $(\overline{e} - \underline{e}) \to 0$ under both EAR and EPD: As the agent becomes infinitely risk-averse, it becomes optimal for him under both types of opaque 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 any deterministic menu are unaffected.

Under EAR, in the limit as $r \to \infty$ and $\sigma^2 = \frac{R}{r} \to 0$, both $\overline{e}$ 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 (6) and (7), simplifies to

$$\lim_{r\to\infty,\sigma^2=R/r\to 0} \Pi^{EAR}(B,k) = \frac{(\delta+1)}{\delta}\frac{B}{(\lambda+1)^2} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2}RB^2\frac{1+2\rho k+k^2}{(1+k)^2}. \tag{13}$$

Under EPD, the principal's payoff for given B and k approaches the same value as under EAR as $r \to \infty$ and $\sigma^2 = \frac{R}{r} \to 0$:[25]

$$\lim_{r\to\infty,\sigma^2=R/r\to 0} \Pi^{EPD}(B,k) = \lim_{r\to\infty,\sigma^2=R/r\to 0} \Pi^{EAR}(B,k). \tag{14}$$

Equations (13) and (14) 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 $(\overline{e} - \underline{e})$ is approximately 0, is to improve the diversification of the shocks to measured performance. Hence, under both EAR and EPD, just as for the case where $\lambda \to 1^+$, it is optimal to set k arbitrarily close to, but less than, $\frac{1}{\lambda}$ ($k \to (\frac{1}{\lambda})^-$). Doing so generates for the principal a payoff approaching

$$\lim_{k\to(1/\lambda)^-}\ \lim_{r\to\infty,\sigma^2=R/r\to 0} \Pi^{EAR}(B,k) = \frac{(\delta+1)}{\delta}\frac{B}{(\lambda+1)^2} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2}RB^2\frac{\lambda^2+2\rho\lambda+1}{(\lambda+1)^2}. \tag{15}$$

The right-hand side of (15) 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.[26]

Thus as $r \to \infty$ and $\sigma^2 = \frac{R}{r} \to 0$, optimally weighted EAR and EPD allow the principal, for any B, to get arbitrarily close to his payoff in the NHI benchmark. Since, by Proposition 1, the best deterministic

[25] This can be proved by first substituting $B = \beta(1 + k)$ into equation (10) and then noting that the terms on the second line all tend to 0 as $r \to \infty$ and $\sigma^2 = \frac{R}{r} \to 0$. The equality in (14) 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 $\sigma^2 \to 0$, $\min\{x_1 + kx_2, x_2 + kx_1\}$ 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 (13).

[26] See equation (24) in the appendix, and set $r\sigma^2 = R$.

menu under hidden information leaves the principal strictly worse off than in the NHI benchmark whenever $\delta > \delta^{NHI}(\lambda, R, \rho)$, we have proved:

Proposition 8 *Consider the limiting case where $r \to \infty$ and $\sigma^2 = \frac{R}{r} \to 0$. Under both EAR and EPD, for any given level of aggregate effort, $\overline{e} + \lambda \underline{e}$, to be induced:*
(i) the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any λ and for any $k < \frac{1}{\lambda}$;
(ii) the optimal value of $k \to \left(\frac{1}{\lambda}\right)^-$;
(iii) with k adjusted optimally, the principal's payoff under both EAR and EPD 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 and β adjusted optimally strictly dominate the best deterministic menu under hidden information.

6.3 Very High Correlation between the Shocks

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

$$\lim_{\rho \to 1} \Pi^{EPD}(B,k) = \frac{\delta+1}{\delta} \frac{B}{(\lambda+1)^2} - \frac{B^2}{2(\lambda+1)^2} - \frac{1}{2} r\sigma^2 B^2. \tag{16}$$

This limiting payoff is independent of k as long as $k < \frac{1}{\lambda}$: As $\rho \to 1$, any possibility of diversifying the risk from the shocks, by increasing k, disappears, and hence the risk cost of EPD becomes $\frac{r\sigma^2 B^2}{2}$. Since (16) is independent of k, any value of k in $(-1, \frac{1}{\lambda})$ is optimal under EPD when $\rho \to 1$. Furthermore, (16) 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 $\beta = \frac{B}{1+\lambda}$.[27] Thus, in this limiting environment as well, EPD yields the principal as high a payoff as in the absence of hidden information, for any level of aggregate effort to be induced.

Under EAR, too, as $\rho \to 1$, diversification of the risk from the shocks becomes impossible, so in equation (6), 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 shows that lowering the weighting parameter k, thereby making the two possible compensation schedules more different, strengthens the agent's incentives for balancing efforts. In the limit as $\rho \to 1$, equations (6) and (7) 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, raises the principal's benefit as well as reducing the risk cost of the exogenous randomization. Hence as $\rho \to 1$, it is optimal under EAR, for any level of aggregate effort to be induced, to set k arbitrarily close to, but larger than, -1 ($k \to -1^+$),

[27] See equation (24) in the appendix, and set $\rho = 1$.

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 (16).[28]

Thus, as $\rho \to 1$, optimally weighted EAR and EPD both allow the principal to approach his payoff in the NHI benchmark, for any B. Combining these results with Proposition 1 yields:

Proposition 9 *Consider the limiting case of perfect correlation of the shocks: $\rho \to 1$. For any given level of aggregate effort, $\overline{e} + \lambda \underline{e}$, to be induced:*
(i) under EPD, the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any λ and for any $k \in (-1, \frac{1}{\lambda})$, and any $k \in (-1, \frac{1}{\lambda})$ is optimal;
(ii) under EAR, the gap in efforts, $\overline{e} - \underline{e}$, approaches 0 for any λ as $k \to -1^+$, and the optimal value of $k \to -1^+$;
(iii) with k adjusted optimally, the principal's payoff under EAR and EPD 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 and β adjusted optimally strictly dominate the best deterministic menu under hidden information.

Analogous arguments and conclusions hold in the limiting environment where the variance σ^2 of the shocks to measured performance approaches 0, holding risk aversion r fixed. 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 these schemes allow the principal to achieve a payoff arbitrarily close to what he would achieve in the absence of *any* informational asymmetries. For $\lambda > 1$, the best deterministic menu under hidden information cannot achieve balanced efforts, even as $\sigma^2 \to 0$, so the best deterministic menu is dominated in this environment by optimally designed EAR and EPD, as long as $\delta > \delta^{NHI}(\lambda, 0, \rho) = \lambda$.

6.4 Discussion

We have identified three environments in which our two simple types of opaque incentive schemes, when designed optimally, strictly dominate the best deterministic menu. Although our propositions focus on limiting environments, our results identify what characteristics of contracting environments increase the relative attractiveness of EAR and EPD. The general message is that, if tasks are sufficiently complementary for the principal (δ sufficiently large), EAR and EPD are superior to deterministic menus in settings where they can provide the agent with very strong incentives for balanced efforts at low cost in terms of the risk imposed. EAR and EPD are more likely to generate a favorable incentive/risk tradeoff when i) the agent's privately known preference between tasks is weak (λ is small), so even a small amount of uncertainty about the weights in the compensation schedule provides a strong impetus for effort balance, or ii) the agent is very risk averse (r is large), so opaque schemes generate a powerful self-insurance motive for effort balance. These two factors increase the attractiveness of EAR and EPD in qualitatively the same way. These schemes are also more likely to be preferred when iii) the errors in measuring performance on the tasks have

[28] As 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.

small variance or large correlation ($\sigma^2(1-\rho)$ is small). 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.

7 Extensions and Robustness

7.1 Opaque Ex Post Linear Schemes vs. Nonlinear Contracts

In our model, gaming takes the form of the agent choosing effort allocations that are excessively (from an efficiency standpoint) sensitive to his private information about his costs. We have shown that ex ante randomization and ex post discretion succeed in mitigating the informed agent's incentives for gaming even under the restriction that ex post, the compensation schedule is linear and separable in the performance measures. The feature shared by EAR and EPD that underlies this finding is their opacity: they both make the agent uncertain ex ante about the incentive coefficients in the linear payment rule. The importance of this ex ante uncertainty for limiting the agent's gaming is highlighted by the contrast between the incentive effects of EAR/EPD and those of deterministic menus of linear contracts: the latter are unable to induce both types of agent to choose positive efforts on both tasks, even when the magnitude of the agent's privately known preference between tasks is arbitrarily small.

We now briefly contrast our two simple classes of opaque ex post linear schemes with nonlinear, nonseparable contracts.

The mechanism by which ex ante randomization alleviates incentives for gaming is one that is familiar and readily comprehensible. The desirability of spreading one's efforts across tasks in order to reduce the riskiness of compensation under ex ante randomization is analogous to strategies for diversifying risk in many commonly encountered 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 contract:

$$w(x_1, x_2) = \alpha - \frac{1}{r}\ln\left\{\frac{1}{2}\exp\left[-r\beta(x_1 + kx_2)\right] + \frac{1}{2}\exp\left[-r\beta(x_2 + kx_1)\right]\right\} \tag{17}$$

Faced with such a contract, how might an agent try to 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 contract, 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 contract 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 contract would, *in theory*, allow the principal to achieve a higher expected payoff than under EAR. But such a conclusion rests on the assumption that an agent would choose the same efforts when faced with the very complicated schedule in (17) as when faced with EAR. We contend that people's intuitive familiarity with the benefits of diversifying risk, coupled with the complexity of (17), make the incentives for balanced efforts considerably more salient under EAR than under (17). As a result, the incentive effects of ex ante randomization are likely to be more consistent and more predictable than those of the contract in (17).[29]

Importantly, in the three environments studied in Section 6, the contract in (17), 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 schedule in (17).*

Consider now ex post discretion, under which the choice between two linear compensation schedules is made strategically by the principal, rather than randomly. Even though EPD achieves the same outcome as if the principal could commit to the nonlinear, nonseparable schedule $w = \min\{\alpha + \beta x_1 + k\beta x_2, \alpha + \beta x_2 + k\beta x_1\}$, EPD requires less commitment power on the principal's part. This has two important 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 contract. In such a model, 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 we highlight, but would also allow the principal to reward, and thereby to encourage, the type of performance which in the circumstances turns out to be appropriate.[30]

7.2 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 analysis of deterministic schemes. We explain here that our key findings are robust to imperfect substitutability of efforts. Specifically, it remains true that i) if tasks are sufficiently complementary for the principal, EAR and EPD are superior to deterministic menus in settings where these opaque schemes generate very strong incentives for balanced efforts, and ii) in such settings, EAR and EPD eliminate the efficiency losses from the agent's hidden information.

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

$$c(\bar{e}, \underline{e}) = \frac{1}{2}\left(\bar{e}^2 + 2s\lambda\bar{e}\underline{e} + \lambda^2\underline{e}^2\right), \tag{18}$$

where the parameter $s \in [0, 1]$ measures the degree of substitutability of efforts. Perfect substitutability

[29] Englmaier, 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. Abeler and Jäger (2013), in a real-effort experiment, compare the behavior of subjects faced with a complex tax scheme with that of subjects faced with a simple one, where the two schemes are designed so that the optimal choices, as well as the incentives around the optimum, are exactly the same. They find that not only are the effort choices of subjects faced with the complex scheme more dispersed, but also that these subjects change their efforts less in response to changes in the scheme.

[30] Scott and Triantis (2006) and Prendergast (1999) develop related arguments.

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 $\delta \to \infty$ in the principal's benefit function, so the efforts on the two tasks are perfect complements for the principal.

In the no hidden information benchmark, the principal will offer each type of agent a contract of the form $w = \alpha + \beta\overline{x} + v\beta\underline{x}$ with $v \geq 1$, where $\overline{x}$ (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 it can be shown that the conclusions of Proposition 1 continue to hold: for any $\lambda > 1$, no menu of deterministic linear contracts can induce both types of agent to choose positive efforts on both tasks, and as a result, the principal's maximized payoff under hidden information is bounded away from that 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 influence, through his efforts, which task is more highly rewarded. Interior optimal efforts under EAR satisfy

$$\frac{\partial c}{\partial \overline{e}} + \frac{\partial c}{\partial \underline{e}} = \beta(1 + k) \tag{19}$$

and

$$\exp\left[r\beta(1 - k)(\overline{e} - \underline{e})\right] = \frac{\frac{c_2}{c_1} - k}{1 - k\frac{c_2}{c_1}}, \tag{20}$$

where $\frac{c_2}{c_1} \equiv \frac{\partial c/\partial \underline{e}}{\partial c/\partial \overline{e}} = \frac{s\lambda\overline{e}+\lambda^2\underline{e}}{\overline{e}+s\lambda\underline{e}}$. Equation (20) is a generalized version of (2) in which the constant λ is replaced by the function $\frac{\partial c/\partial \underline{e}}{\partial c/\partial \overline{e}}$. Optimal efforts under EPD satisfy (19) and equation (9), except that the right-hand side of (9) is replaced by the right-hand side of (20).

Consider now the three environments studied in detail in Section 6. As $\lambda \to 1^+$ or as $r \to \infty$, $\sigma^2 \to 0$, it follows from (20) and its analog for EPD that both EAR and EPD induce perfectly balanced efforts for any $k \in (-1, \frac{c_1}{c_2})$.[31] Therefore, in these limiting cases, the only effect of increasing k is to improve the diversification of the risk from the shocks. Hence it is optimal in both environments, with both EAR and EPD, to set k as large as possible subject to keeping efforts perfectly balanced, i.e., to take $k \to (\frac{c_1}{c_2})^-$. Since with perfectly balanced efforts, $\frac{c_1}{c_2} = \frac{1+s\lambda}{\lambda(\lambda+s)} = 1/v^{NHI}$, it follows that as $\lambda \to 1^+$ or as $r \to \infty$, $\sigma^2 \to 0$, 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 $\rho \to 1$, the weight k has no effect on diversification, so it is optimal under EAR and EPD to set k to induce perfectly balanced efforts; in this setting, too, optimally weighted EAR and EPD generate a payoff arbitrarily close to that in the NHI benchmark.

As long as $s \geq \frac{1}{\lambda}$, we saw above that under hidden information, the principal's maximized payoff

[31] If $k > \frac{c_1}{c_2}$, (20) and its analog for EPD show that neither EAR nor EPD can induce interior solutions for efforts.

is bounded away from that in the NHI benchmark, because no linear deterministic menu can induce both types of agent to work on both tasks. Since in the environments studied in Section 6, optimally designed EAR and EPD generate a payoff arbitrarily close to that in the NHI benchmark, it follows that in these settings, these opaque schemes are superior to linear deterministic menus. Hence, allowing the agent's efforts on the tasks to be less than perfect substitutes in his cost function does not alter our main results.

7.3 Beyond the Exponential-Normal Model

Our findings that opaque 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 $U(w - c(\overline{e}, \underline{e}))$, with $U(\cdot)$ an arbitrary strictly concave function and $c(\overline{e}, \underline{e})$, as in (18), reflecting imperfect substitutability of efforts.

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

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

The second equation, which is a generalized version of (5), shows that just as for the exponential-normal model, both EAR and EPD give the risk-averse agent an incentive to choose more balanced efforts to partially self-insure against the risk stemming from the uncertainty about which payment schedule will ultimately be used. Analogously to the exponential-normal case, EPD also provides the agent with a second motive for balancing efforts, because his efforts actually influence which compensation schedule the principal chooses. In this more general setting, however, the strength of the agent's self-insurance motive is not generally equal under the two types of opaque schemes, so EPD does not necessarily induce more balanced efforts than EAR.

Nevertheless, we can show that 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. Moreover, when efforts are perfect substitutes for the agent ($s = 1$), as λ increases from 1, $\overline{e}^{EAR}/\underline{e}^{EAR}$ and $\overline{e}^{EPD}/\underline{e}^{EPD}$ both increase continuously from 1, whereas $\overline{e}^{SD}/\underline{e}^{SD}$ jumps from 1 to ∞. Thus, even outside the exponential-normal framework, the opaque schemes provide stronger incentives for effort balance and are more robust to hidden information.

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

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

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; the lump-sum payment is always equal to 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. Details of the derivations are in Appendix B.

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

$$\lambda\underline{e} + (n-1)\overline{e} = \frac{\kappa\beta}{n-1+\lambda} \qquad \text{and} \qquad \overline{e} - \underline{e} = \frac{1}{r\beta}\ln\left[\frac{\lambda(n-\kappa)}{(n-1)-(\kappa-1)\lambda}\right]. \tag{21}$$

Reducing κ, the number of tasks rewarded, makes the risk imposed by the randomization more costly, so strengthens the agent's incentive to self-insure. As a result, the agent's optimal effort profile is more balanced ($\overline{e} - \underline{e}$ is smaller), the smaller is the number of tasks rewarded.

Since aggregate effort is proportional to $\kappa\beta$, define $\tilde{\beta} \equiv \kappa\beta$. Using (21), we can write the principal's payoff as a function of $\tilde{\beta}$ and κ, when α is set to ensure zero rent for each type of agent:

$$\begin{aligned}\Pi(\tilde{\beta}, \kappa) = \underline{e} + \frac{(n-1)}{\delta}\overline{e} - &\frac{\tilde{\beta}^2}{2\,(n-1+\lambda)^2} \\ &- \frac{1}{2}r\sigma^2\tilde{\beta}^2\frac{1+\rho(\kappa-1)}{\kappa} - \frac{1}{nr}\ln\left[\frac{(n-\kappa)^{n-\kappa}(n-1+\lambda)^n}{n^n\lambda^\kappa\left((n-1)-(\kappa-1)\lambda\right)^{n-\kappa}}\right],\end{aligned} \tag{22}$$

where

$$\underline{e} + \frac{n-1}{\delta}\overline{e} = \frac{\delta+n-1}{\delta}\frac{\tilde{\beta}}{(n-1+\lambda)^2} - \frac{(\delta-\lambda)(n-1)\kappa}{\delta(n-1+\lambda)r\tilde{\beta}}\ln\left[\frac{\lambda(n-\kappa)}{(n-1)-(\kappa-1)\lambda}\right]. \tag{23}$$

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 (22)-(23) with equations (6)-(7) reveals that changes in κ have qualitatively the same three effects on the principal's payoff in this n-task model as do variations in the weighting coefficient k in EAR in the original two-task model. Specifically, an increase in κ, by inducing a larger gap $\overline{e} - \underline{e}$, has two negative effects: i) it lowers the principal's benefit $\underline{e} + \frac{n-1}{\delta}\overline{e}$ when aggregate effort is held fixed, as long as $\delta > \lambda$ (this corresponds to the fact that (23) is decreasing in κ), and ii) it raises the cost of compensating the agent for the risk imposed by the exogenous randomization (this corresponds to the fact that the term in square brackets in (22) is increasing in κ). At the same time, raising κ also improves the diversification of the risk from the shocks to measured performance (as reflected in the fact that $\frac{1+\rho(\kappa-1)}{\kappa}$ in (22) 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 to derive the following comparative statics results for the optimal number of tasks to reward, given any desired level of aggregate effort. Analogously with Proposition 3, the optimal number of tasks to reward is smaller, i) the stronger is the principal's preference for balanced efforts (i.e., the larger is δ); ii) the less risk-averse the agent is, holding $r\sigma^2$ fixed; iii) the lower the importance of diversification of the risk from the shocks to measured performance (i.e., the lower is $\sigma^2(1-\rho)$); and the smaller the level of aggregate effort to be induced.

7.5 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 $p \to 1/2$, 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 $\overline{e} - \underline{e}$ 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 opaque 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.

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, and US law school rankings, among others. 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 "opacity" about the criteria upon which rewards will be based and/or how heavily these criteria will be weighted–has been used or advocated to deter gaming.

To elucidate the benefits and costs of opaque incentive schemes, we have analyzed a model in which the agent is better informed than the principal about his costs of effort on different tasks. The agent games transparent schemes by choosing effort allocations that are excessively (from an efficiency perspective) sensitive to his private information. We studied two simple classes of opaque schemes, ex ante randomization (EAR) and ex post discretion (EPD), each of which makes the agent uncertain ex ante about the incentive coefficients in the linear payment rule. We explored how and to what extent each of these opaque schemes mitigates the agent's gaming. When the agent has private information about his costs of effort on different tasks, the principal in general faces a trade-off between the benefits of the more efficient effort allocations induced by opaque schemes and the costs of the greater risk they impose.

Our key contribution has been, in Propositions 7, 8, and 9, to identify environments in which optimally designed opaque schemes strictly outperform all linear deterministic menus. In each of these environments, optimally weighted ex ante randomization and ex post discretion induce the agent to choose the socially efficient, perfectly balanced effort profile, 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 opaque schemes can strictly dominate all deterministic menus, our results identify what characteristics of contracting environments increase the relative attractiveness of opaque schemes. Ex ante randomization and ex post discretion are more likely to be preferred when i) efforts on the tasks are highly complementary for the principal; ii) the agent's privately known preference between tasks is weak, so even a small amount of uncertainty about the weights in the compensation schedule induces a relatively balanced effort profile; and iii) the agent is very risk averse, so opaque schemes generate a powerful self-insurance motive for balancing efforts. These three factors increase the attractiveness of EAR and EPD in qualitatively the same way. These schemes are also more likely to be preferred when iv) the errors in measuring performance on the tasks have small variance or large correlation. 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, opaque 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 clearly significantly enhanced relative to the best deterministic contract 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 would 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.

Our analysis suggests that even beyond the specific multi-task setting on which we have focused, opacity of incentive schemes, by making agents more uncertain about the consequences of their actions for their rewards, could help principals to mitigate the costs of gaming stemming from agents' exploiting their better knowledge of the environment. Future research should explore the benefits of opaque incentive schemes in deterring gaming in other settings, seeking to identify under what conditions these incentive benefits can outweigh the risk costs of opacity.

References

[1] **Abeler, Johannes and Simon Jäger**, "Complex Tax Incentives–An Experimental Investigation," *CESifo Working Paper No. 4231*, 2013.

[2] **Baker, George P.**, "Incentive Contracts and Performance Measurement", *Journal of Political Economy*, Jun. 1992, *100* (3), 598-614.

[3] **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.

[4] **Baliga, Sandeep and Tomas Sjöström**, "Strategic Ambiguity and Arms Proliferation," *Journal of Political Economy*, Dec. 2008, *116* (6), 1023-1057.

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

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

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

[8] **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.

[9] **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.

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

[11] **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.

[12] **Carroll, Gabriel**, "Robustness and Linear Contracts," *Working Paper*, 2012.

[13] **Chetty, Raj, Looney, Adam and Kory Kroft**, "Salience and Taxation: Theory and Evidence," *American Economic Review*, Sep. 2009, *99* (4), 1145-1177.

[14] **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.

[15] **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.

[16] **Forbes, Silke, Lederman, Mara and Trevor Tombe**, "Quality Disclosure Programs with Thresholds: Misreporting, Gaming and Employee Incentives," *Working Paper*, 2012.

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

[18] **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.

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

[20] **Healy, Andrew and Jennifer Pate**, "Can Teams Help to Close the Gender Competition Gap?" *Economic Journal*, Sep. 2011, *121* (6), 1192-1204.

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

[22] — **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.

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

[24] **Jehiel, Philippe and Andrew F. Newman**, "Loopholes: Social Learning and the Evolution of Contract Form, " *Working Paper*, 2011.

[25] **Larkin, Ian**, "The Cost of High-Powered Incentives: Employee Gaming in Enterprise Software Sales," *Journal of Labor Economics*, forthcoming.

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

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

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

[29] **Milgrom, Paul R. and D. John Roberts**, *Economics, Organization and Management*, Englewood Cliffs: Prentice-Hall, 1992.

[30] **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.

[31] **Osler, Mark,** "The Case Against Transparency in Law School Rankings Methodology," *Law School Innovation*, May 4, 2010, http://lsi.typepad.com/lsi/2010/05/the-case-against-transparency-in-law-school-rankings-methodology.html.

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

[33] **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.

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

[35] **Rahman, David**, "But Who Will Monitor the Monitor?" *American Economic Review*, Oct. 2012, 102 (6), 2767-2797.

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

[37] **Stiglitz, Joseph**, "Self-selection and Pareto Efficient Taxation," *Journal of Public Economics*, Mar. 1982, *17* (2), 213–240.

[38] **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. 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

$$\Pi^{bal}(\beta) = \frac{\beta}{1+\lambda}\left(1 + \frac{1}{\delta}\right) - \frac{\beta^2}{2} - \frac{1}{2} r\sigma^2 \beta^2 (1 + 2\rho\lambda + \lambda^2). \tag{24}$$

With β chosen optimally, the resulting maximized payoff is

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

This payoff is continuous as $\lambda \to 1$.

Now consider the pair of contracts (C_1^{foc}, C_2^{foc}). C_i^{foc} induces type i to choose $e_i = \beta$ and $e_j = 0$. 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 then be

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

With β chosen optimally, the resulting maximized payoff is

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

Comparison of the expressions for Π^{bal} and Π^{foc} shows that there is a critical value of δ,

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

above (below) which $\Pi^{bal} > (<)\, \Pi^{foc}$. It is straightforward to verify that δ^{NHI} is increasing in each of its arguments. ■

Proof of Proposition 1. To prove part (i), observe that when faced with a menu of deterministic linear contracts, an agent either is willing to exert perfectly balanced efforts or strictly prefers fully focused efforts. Therefore, if a menu existed which could induce both types to choose strictly positive efforts on both tasks, it would have the form

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

and would induce agent i to choose C_i. 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 strictly positive efforts on both tasks.

Parts (ii) and (iii) are proved in the text following the statement of the proposition. ■

Proof of Proposition 2.

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

$$\frac{1}{2}\overline{EU} + \frac{1}{2}\underline{EU} = -\frac{1}{2}E\exp\left\{-r\left[\alpha + \beta\overline{x} + k\beta\underline{x} - \frac{1}{2}(\overline{e}+\lambda\underline{e})^2\right]\right\} - \frac{1}{2}E\exp\left\{-r\left[\alpha + \beta\underline{x} + k\beta\overline{x} - \frac{1}{2}(\overline{e}+\lambda\underline{e})^2\right]\right\}$$

$$= -\frac{1}{2}\exp\left\{-r\left[\alpha + \beta\overline{e} + k\beta\underline{e} - \frac{r}{2}\sigma^2\beta^2(1+2\rho k+k^2) - \frac{1}{2}(\overline{e}+\lambda\underline{e})^2\right]\right\}$$
$$- \frac{1}{2}\exp\left\{-r\left[\alpha + \beta\underline{e} + k\beta\overline{e} - \frac{r}{2}\sigma^2\beta^2(1+2\rho k+k^2) - \frac{1}{2}(\overline{e}+\lambda\underline{e})^2\right]\right\} \quad (26)$$

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

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

These first-order conditions can be rewritten as

$$\beta\overline{EU} + k\beta\underline{EU} = (\overline{e}+\lambda\underline{e})(\overline{EU}+\underline{EU}) \quad (27)$$
$$k\beta\overline{EU} + \beta\underline{EU} = \lambda(\overline{e}+\lambda\underline{e})(\overline{EU}+\underline{EU}). \quad (28)$$

Equations (27) and (28) 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 (27) and (28) and rearranging yields equation (1). Using (1) to substitute for aggregate effort $(\overline{e}+\lambda\underline{e})$ in (27) 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 $\overline{e} - \underline{e}$ yields $\overline{e} - \underline{e} = [\ln(\frac{\lambda-k}{1-k\lambda})]/[r\beta(1-k)]$. For $k \in \left(-1, \frac{1}{\lambda}\right)$ and $\lambda > 1$, therefore, $(\overline{e} - \underline{e})$ is greater than 0, increasing in λ and k, and decreasing in r. $(\overline{e} - \underline{e}) \to 0$ as $\lambda \to 1$, $k \to -1^+$, or $r \to \infty$.

Proof of Part (iv): Using (1) and (2) to substitute into (26), and then simplifying, allows us to express each type of agent's expected utility under EAR as

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

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:

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

where the final line uses (2). ■

Proof of Proposition 3. Define $B \equiv \beta(1+k)$ and note, from (1), that aggregate effort $\overline{e}+\lambda\underline{e}$ is proportional to B. Using (1), (2), and $\beta = \frac{B}{1+k}$ to substitute into (3) yields (6) and (7) 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 (6) and (7) 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 (7) 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 (7) and fourth term in (6) 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 (6), and $\frac{\partial \Pi}{\partial k}$ is increasing in $\sigma^2(1-\rho)$, so the optimal k decreases as $\sigma^2(1-\rho)$ decreases.
Part (iv): $\frac{\partial^2 \Pi}{\partial B \partial k} > 0$, 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 $\overline{e}$, $\underline{e}$, $\overline{x}$, 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 $w = \alpha + \beta \min\{\overline{x} + k\underline{x}, \underline{x} + k\overline{x}\}$, each type's expected utility is

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

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

$$\left(\sigma^k\right)^2 \equiv var\left(x_1 + kx_2\right) = \sigma^2\left(1 + 2\rho k + k^2\right) \quad \text{and} \quad \rho^k \equiv corr\left(x_1 + kx_2, x_2 + kx_1\right) = \frac{\rho + 2k + \rho k^2}{1 + 2\rho k + k^2}. \tag{30}$$

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

$$\theta \equiv \sigma\left[2\left(1-\rho\right)\right]^{\frac{1}{2}} \quad \text{and} \quad \theta^k \equiv \sigma^k\left[2\left(1-\rho^k\right)\right]^{\frac{1}{2}} = \theta(1-k). \tag{31}$$

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

$$\begin{aligned} m\left(-r\beta\right) \quad \equiv \quad & E\left\{\exp\left[-r\beta \min\left\{\overline{x} + k\underline{x}, \underline{x} + k\overline{x}\right\}\right]\right\} \\ & + \exp\left[-r\beta\left(\underline{e} + k\overline{e}\right) + \frac{1}{2}r^2\beta^2\left(\sigma^k\right)^2\right]\Phi\left(\frac{\left(\overline{e} - \underline{e}\right)\left(1-k\right) + r\beta\left(\sigma^k\right)^2\left(1-\rho^k\right)}{\theta^k}\right) \\ = \quad & \exp\left[-r\beta\left(\overline{e} + k\underline{e}\right) + \frac{1}{2}r^2\beta^2\left(\sigma^k\right)^2\right]\Phi\left(\frac{-(\overline{e} - \underline{e})}{\theta} + \frac{r\beta\theta(1-k)}{2}\right) \\ & + \exp\left[-r\beta\left(\underline{e} + k\overline{e}\right) + \frac{1}{2}r^2\beta^2\left(\sigma^k\right)^2\right]\Phi\left(\frac{(\overline{e} - \underline{e})}{\theta} + \frac{r\beta\theta(1-k)}{2}\right) \\ = \quad & \exp\left[-r\beta\left(\overline{e} + k\underline{e}\right) + \frac{1}{2}r^2\beta^2\left(\sigma^k\right)^2\right]\Phi(-) + \exp\left[-r\beta\left(\underline{e} + k\overline{e}\right) + \frac{1}{2}r^2\beta^2\left(\sigma^k\right)^2\right]\Phi(+). \end{aligned} \tag{32}$$

Using (32) in (29), we can derive and simplify the first-order conditions for interior optimal values of $\overline{e}$ and $\underline{e}$, respectively:

$$(\overline{e} + \lambda\underline{e})m(-r\beta) = \beta \exp\left[-r\beta\left(\overline{e} + k\underline{e}\right) + \frac{1}{2}r^2\beta^2\left(\sigma^k\right)^2\right]\Phi(-) + k\beta \exp\left[-r\beta\left(\underline{e} + k\overline{e}\right) + \frac{1}{2}r^2\beta^2\left(\sigma^k\right)^2\right]\Phi(+) \tag{33}$$

$$\lambda(\overline{e} + \lambda\underline{e})m(-r\beta) = k\beta \exp\left[-r\beta\left(\overline{e} + k\underline{e}\right) + \frac{1}{2}r^2\beta^2\left(\sigma^k\right)^2\right]\Phi(-) + \beta \exp\left[-r\beta\left(\underline{e} + k\overline{e}\right) + \frac{1}{2}r^2\beta^2\left(\sigma^k\right)^2\right]\Phi(+). \tag{34}$$

Equations (33) and (34) imply

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

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 (33) and (34) and rearranging yields equation (8). Using (8) to substitute for aggregate effort $(\overline{e} + \lambda \underline{e})$ in (33) yields, after simplification, equation (9).

Proof of Part (iii): Define $d \equiv \overline{e} - \underline{e}$ and rewrite (9) in terms of d as

$$\exp\left[r\beta\left(1-k\right)d\right] \frac{\Phi\left(\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)}{\Phi\left(-\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)} = \frac{\lambda - k}{1 - k\lambda}. \tag{35}$$

The left-hand side of (35) is increasing in d and equals 1 at $d = 0$. The right-hand side is increasing in λ and $\to 1$ as $\lambda \to 1$, so d is increasing in λ and $d \to 0$ as $\lambda \to 1$. Differentiating (35) implicitly w.r.t. k yields

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

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

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

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

$$y \equiv \frac{d}{\theta} \qquad \text{and} \qquad t \equiv \frac{r\beta\theta(1-k)}{2}. \tag{36}$$

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

$$\frac{2d}{\theta} + \frac{\phi(+)}{\Phi(+)} - \frac{\phi(-)}{\Phi(-)} = y + t + \frac{\phi(y+t)}{\Phi(y+t)} - \left(-y + t + \frac{\phi(-y+t)}{\Phi(-y+t)}\right) = j(y+t) - j(-y+t), \tag{37}$$

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 $k \to -1^+$, the right-hand side of (35) approaches 1 , so d approaches 0.

Differentiating (35) implicitly with respect to r and rearranging shows that $\frac{\partial d}{\partial r}$ has the opposite sign to the expressions in (37), so d is decreasing in r. As $r \to \infty$, the ratio $\frac{\Phi(+)}{\Phi(-)}$ on the left-hand side of (35) approaches 1, so for the left-hand side as a whole to remain finite requires $d \to 0$. Hence as $r \to \infty$, $d \to 0$.

To prove that d is increasing in $\sigma^2(1-\rho) = \frac{\theta^2}{2}$, differentiate (35) 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, $d \to 0$, since otherwise the left-hand side of (35) becomes infinite.

Proof of Part (iv): Using (8), (9), and (32) to substitute into (29) allows us to express each type of agent's expected utility under EPD as

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

Hence each type of agent's certainty equivalent is

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

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

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

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

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

which is the payoff expression given in (10).

Proof of Part (v): The left-hand sides of (2) and (9) are both increasing in $d \equiv (\overline{e} - \underline{e})$, and since $\Phi(+) > \Phi(-)$ whenever $d > 0$, the left-hand side of (9) 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 (10) 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:

$$\Pi^{EPD}(\beta,k) - \left[\underline{e}^{EPD} + \frac{1}{\delta}\overline{e}^{EPD} - \frac{\beta^2(1+k)^2}{2(\lambda+1)^2} - \frac{1}{2}r(\sigma)^2\beta^2\left(1+2\rho k+k^2\right)\right] \geq 0, \tag{38}$$

(38) 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 (38), we use (10) to express the left-hand side as

$$\begin{aligned}
\frac{1}{r}\Bigg\{&-r\beta d(1-k)\Phi\left(\frac{-d}{\theta}\right) + r\beta\theta(1-k)\phi\left(\frac{d}{\theta}\right) \\
&- \ln\left[\exp\{-r\beta(1-k)d\}\Phi\left(\frac{-d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right) + \Phi\left(\frac{d}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)\right]\Bigg\}.
\end{aligned} \tag{39}$$

The terms in curly brackets in (39) can be rewritten as

$$h(y,t) \equiv -2ty\Phi(-y) + 2t\phi(-y) - \ln\left[\exp\{-2ty\}\Phi(-y+t) + \Phi(y+t)\right], \tag{40}$$

where we have used the definitions of y and t in (36) 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 $y \to \infty$, $h(y,t) \to 0$ 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 $\frac{\partial h(y,t)}{\partial y} \le 0$. Now

$$\frac{\partial h(y,t)}{\partial y} \overset{\text{sgn}}{=} -\Phi(-y) + \frac{\Phi(-y+t)}{\exp\{2ty\}\Phi(y+t) + \Phi(-y+t)}. \tag{41}$$

Define $q(y,t)$ to equal the right-hand side of (41). $q(y,0) = 0$, so showing that $q(y,t)$ is decreasing in t for all $t \ge 0$ will imply that $\frac{\partial h(y,t)}{\partial y} \le 0$.

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

since $j(z) = z + \frac{\phi(z)}{\Phi(z)}$ is increasing (see (37)). Hence for all $y \ge 0$, $t \ge 0$, $\frac{\partial h(y,t)}{\partial y} \le 0$, and so $h(y,t) \ge 0$.
Step 2: When $\delta \ge \lambda$,

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

This step follows from the facts that aggregate effort $\overline{e} + \lambda\underline{e}$ is equal under EPD and EAR (as shown by (1) and (8)) and that the gap in efforts, $\overline{e} - \underline{e}$, is smaller under EPD than EAR.
Step 3:

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

This step follows from (3), $\lambda \ge 1$, 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 $\overline{e}^{EPD} + \lambda\underline{e}^{EPD} = \frac{\beta(1+k)}{2}$, and $\overline{e}^{EPD} = \underline{e}^{EPD} = \frac{\beta(1+k)}{4}$. Hence, for $\lambda = 1$, equation (10) simplifies to

$$\Pi^{EPD}(\beta,k) = \frac{\delta+1}{\delta}\frac{\beta(1+k)}{4} - \frac{1}{8}\beta^2(1+k)^2 - \frac{1}{2}r(\sigma^k)^2\beta^2 - \frac{1}{r}\left\{\ln\left[2\Phi\left(\frac{r\beta\theta(1-k)}{2}\right)\right] - r\beta\theta(1-k)\phi(0)\right\}. \tag{42}$$

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

$$\Pi^{SD}(\beta^{SD}) = \frac{\delta+1}{\delta}\frac{\beta^{SD}}{2} - \frac{1}{2}\left(\beta^{SD}\right)^2 - r\sigma^2\left(\beta^{SD}\right)^2(1+\rho) = \frac{\delta+1}{\delta}\frac{\beta(1+k)}{4} - \frac{1}{8}\beta^2(1+k)^2 - \frac{1}{4}r\sigma^2\beta^2(1+k)^2(1+\rho). \tag{43}$$

Using (42) and (43) and the definition of $(\sigma^k)^2$ in (30), we have

$$\begin{aligned}\Pi^{SD}(\beta^{SD}) - \Pi^{EPD}(\beta,k) &= \frac{r\beta^2}{2}\left[(\sigma^k)^2 - \frac{\sigma^2(1+k)^2)(1+\rho)}{2}\right] + \frac{1}{r}\left\{\ln\left[2\Phi\left(\frac{r\beta\theta(1-k)}{2}\right)\right] - r\beta\theta(1-k)\phi(0)\right\} \\ &= \frac{1}{4}r\sigma^2\beta^2(1-\rho)(1-k)^2 + \frac{1}{r}\left\{\ln\left[2\Phi\left(\frac{r\beta\theta(1-k)}{2}\right)\right] - r\beta\theta(1-k)\phi(0)\right\}. \end{aligned} \tag{44}$$

(44) has the sign of $g(t) \equiv \frac{t^2}{2} + \ln[2\Phi(t)] - 2t\phi(0)$, where t is defined in (36). We have $g(0) = 0$. Also,

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

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 $\underline{e}^{EPD} = 0$), the first-order condition (33) for $\bar{e}^{EPD}$ reduces to (contrast this with (9)):

$$\exp\left\{r\beta(1-k)\bar{e}^{EPD}\right\} \frac{\Phi\left(\frac{\bar{e}^{EPD}}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)}{\Phi\left(\frac{-\bar{e}^{EPD}}{\theta} + \frac{r\beta\theta(1-k)}{2}\right)} = \frac{\beta - \bar{e}^{EPD}}{\bar{e}^{EPD} - k\beta}. \tag{45}$$

Since the left-hand side of (45) is strictly greater than 1 for $k < 1$, (45) implies that $\bar{e}^{EPD} < \frac{\beta(1+k)}{2}$. When EPD induces A to choose the corner solution $\left(\bar{e}^{EPD}, 0\right)$,

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

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

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

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

$$\begin{aligned}\frac{r^2\sigma^2}{4}[2\beta^2(1+2\rho k+k^2) - 4(1+\rho)\left(\bar{e}^{EPD}\right)^2] \\ + \ln\left[\Phi(+) + \exp\left\{-r\beta(1-k)\bar{e}^{EPD}\right\}\Phi(-)\right] \\ + r\beta(1-k)\bar{e}^{EPD}\Phi\left(\frac{-\bar{e}^{EPD}}{\theta}\right) + r\beta\theta(1-k)\phi\left(\frac{\bar{e}^{EPD}}{\theta}\right).\end{aligned} \tag{46}$$

Since (45) implies that $\bar{e}^{EPD} < \frac{\beta(1+k)}{2}$, the expression on the first line of (46) is strictly greater than

$$\frac{r^2\beta^2\sigma^2}{4}\left[2(1+2\rho k+k^2) - (1+\rho)(1+k)^2\right]. \tag{47}$$

Using (47), the definitions of y and t in (36), and the second and third lines of (46), we conclude that

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

where $h(y,t)$ is defined in (40). We showed there that $h(y,t)$ is decreasing in y, for all $y \geq 0$, $t \geq 0$, hence

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

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

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 $\underline{e}^{EAR} = 0$ (compare (2)):

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

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

$$\begin{aligned}
\Pi^{EPD}(\beta,k) &= \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}\left(\bar{e} + \lambda\underline{e}\right)^2 - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}h(y,t) \\
&\leq \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}\left(\bar{e} + \lambda\underline{e}\right)^2 - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}h(0,t) \\
&= \underline{e} + \frac{\bar{e}}{\delta} - \frac{1}{2}\left(\bar{e} + \lambda\underline{e}\right)^2 - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}[-\ln(2\Phi(t)) + 2t\phi(0)] \\
&< \frac{\bar{e} + \lambda\underline{e}}{\delta} - \frac{1}{2}\left(\bar{e} + \lambda\underline{e}\right)^2 - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}\left[-\ln\left(2\Phi\left(t\right)\right) + 2t\phi\left(0\right)\right] \\
&= \frac{\beta(1+k)}{\delta\left(1+\lambda\right)} - \frac{1}{2}\frac{\beta^2(1+k)^2}{\left(1+\lambda\right)^2} - \frac{1}{2}r\beta^2(\sigma^k)^2 + \frac{1}{r}\left[-\ln\left(2\Phi\left(t\right)\right) + 2t\phi\left(0\right)\right], \qquad (48)
\end{aligned}$$

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 (8).

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

$$\Pi^{SD}(\beta^{SD}) = \frac{\beta(1+k)}{\delta\left(1+\lambda\right)} - \frac{1}{2}\frac{\beta^2(1+k)^2}{\left(1+\lambda\right)^2} - r\sigma^2\beta^2(1+k)^2\frac{(1+\rho)}{(1+\lambda)^2}. \qquad (49)$$

Hence from (48) and (49) we can conclude that

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

where the strict inequality is a consequence of the inequalities in (48), the weak inequality follows since $\lambda \geq 1$, the first equality uses (30), the second equality uses (36), 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 7.4 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 $\overline{e}$, 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, $\overline{EU}$). 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

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

We focus on the case where optimal efforts are interior.

The aggregate effort exerted by an agent is $\lambda\underline{e} + (n-1)\,\overline{e}$, 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 $\overline{e} - \underline{e}$, we need the first-order condition for $\underline{e}$, which is

$$\binom{n-1}{\kappa-1}\left[\beta - \lambda A\right]\underline{EU} + \binom{n-1}{\kappa}\left[-\lambda A\right]\overline{EU} = 0, \tag{50}$$

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 (50) and rearranging yields

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

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

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

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 each type of agent's optimal choices of A and $\overline{e} - \underline{e}$, the principal's expected payoff as a function of β and κ can be simplified to

$$\begin{aligned} \Pi(\beta,\kappa) = &\left(\underline{e} + \frac{(n-1)}{\delta}\overline{e}\right) - \frac{\kappa^2\beta^2}{2\,(n-1+\lambda)^2} \\ &- \frac{1}{2}r\sigma^2\beta^2\kappa\left(1+\rho(\kappa-1)\right) - \frac{1}{nr}\ln\left[\frac{(n-\kappa)^{n-\kappa}(n-1+\lambda)^n}{n^n\lambda^\kappa\left((n-1)-(\kappa-1)\lambda\right)^{n-\kappa}}\right], \end{aligned} \tag{51}$$

where

$$\underline{e} + \frac{(n-1)}{\delta}\overline{e} = \left(\frac{\delta+n-1}{\delta}\right)\frac{\kappa\beta}{(n-1+\lambda)^2} - \frac{(\delta-\lambda)(n-1)}{\delta(n-1+\lambda)r\beta}\ln\left[\frac{\lambda(n-\kappa)}{(n-1)-(\kappa-1)\lambda}\right]. \tag{52}$$

Using $\tilde{\beta} = \kappa\beta$ to substitute for β in the above payoff expression yields expressions (22) and (23) in the text.

Verifying the claims in the final paragraph of Section 7.4 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 (22) 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 $\frac{\partial^2 \Pi}{\partial \tilde{\beta} \partial \kappa} > 0$.