Contracting with Private Rewards*

René Kirkegaard†

February 2015

Abstract

I extend the basic moral hazard model to include the possibility that the agent faces non-contractible uncertainty that is payoff relevant to him. These rewards may or may not be of a monetary nature. The agent devotes effort towards working for the principal and towards pursuing private rewards. Thus, the agent is multi-tasking. Despite the addition of private rewards and multi-tasking, I establish conditions under which the first-order approach remains valid. Interestingly, the model identifies a new source of economic rents for the agent. Moreover, to induce higher effort on the job, it may be optimal to write a contract that to an outsider appears to offer higher base utility but weaker or flatter incentives on the margin. Thus, the paper contributes to the literature on extrinsic versus intrinsic motivation. Here, the explanation is that a contract change leads the agent to reevaluate his "work-life balance" and to adjust, jointly, the effort devoted to both types of rewards. The principal optimally induces higher effort by manipulating this balance through the contract design. Finally, implications for common agency are examined.

JEL Classification Numbers: D82, D86

Keywords: Common Agency, First-Order Approach, Moral Hazard, Multi-tasking, Principal-Agent Models, Private Rewards.

*I thank the Canada Research Chairs programme and SSHRC for funding this research. I am grateful for comments and suggestions from a seminar audience at Queen's University.

†Department of Economics and Finance, University of Guelph. Email: rkirkega@uoguelph.ca.

1 Introduction

The principal-agent model has been tremendously influential in economics. However, the canonical model essentially assumes that the principal-agent relationship takes place in a perfect vacuum – there are no (non-contractible) outside random disturbances. For instance, the only payoff-relevant risk the agent faces is due to the uncertainty embodied in the incentive scheme presented to him by the principal.

In reality, however, it is easier to think of examples in which the agent faces some non-contractible "background risk" than examples in which this is not the case. Indeed, such uncertainty is often endogenous. That is, the agent pursues a host of potentially rewarding activities that are not directly observable (nor necessarily directly relevant) to the principal. Even seemingly mundane activities may in reality entail significant rewards. For instance, when the busy young professional tolerates dinner with her parents, she may hope to become part of the "27 percent of those purchasing a home for the first time [who] received a cash gift from relatives or friends to come up with a down payment."[1] When her older brother moves his family closer to their parents at the cost of a longer commute, he may be motivated by the knowledge that "by the time the average youngster reaches school age, they will have been babysat by their grandparents for more than 5,610 hours."[2] The "rewards" the siblings obtain from their parents are most likely not observable to their employers. In other words, these rewards are non-contractible.

There are a plethora of other examples in which the agent directly receives a reward from a third party. Although the waiter has an employment contract with the restaurant owner, a significant part of her income often comes in the form of tips from the diner, despite the fact that there is no explicit contract between the two (nor is there an explicit contract between the parents and offspring in the previous paragraph).[3] In other cases, the agent is in a formal contractual relationship with more than one principal, a situation known as common agency. This holds for any

[1] The data is for the U.S, in 2013. See www.bloomberg.com/news/2014-09-19/mom-and-dad-banks-step-up-aid-to-first-time-home-buyers.html

[2] The data is for the U.K. The estimated monetary value of this amount of child care is £21,654.60. See http://www.dailymail.co.uk/femail/article-2263843/The-21-000-grandma-Grandparents-babysitting-duties-reduce-cost-childcare-whopping-4-300-EVERY-YEAR.html

[3] The Globe and Mail reports that following a tax audit among 145 wait staff in St. Catharines, ON, Canada, it was estimated that "tip income typically amounted to between 100 and 200 per cent of wage income." See http://www.theglobeandmail.com/report-on-business/economy/economy-lab/tipping-point-ottawa-loses-billions-in-undeclared-income/article4418504/

worker with more than one job. Likewise, any employed home-owner also finds himself in this situation. The employer is one principal, his insurance company another. It is commonly the case that any given principal is unable to observe how well the agent performs on behalf of the other principal. Thus, developing an understanding of contracting with private rewards is a necessary first step towards analyzing such limited-information common agency environments.

Examples involving potentially large non-monetary rewards include the agent's health status as impacted by endogenous life-style choices, his social status in his peer group, the quality of his match on the marriage market as affected by his search intensity, and so on. The satisfaction from mastering a second language, or any other hobby, is another example. Even the agent's "job-satisfaction" may be endogenous, influenced by e.g. the enthusiasm with which he interacts with colleagues. Note that at the time of contracting, the agent may still face uncertainty over his private rewards. As private rewards and labor income are likely to interact in his utility function, the agent is thus uncertain even of his own marginal utility of labor income when he is offered the contract.

The aim of this paper is to analyze the consequences of endogenous "private rewards" on optimal contracting. The standard principal-agent model is amended to allow the agent to work on two tasks. The first "task" captures the effort the agent expends working on behalf of the principal. This task produces a contractible signal, as in the standard model. The second task describes the effort devoted to pursuing private rewards (which the principal may or may not directly care about). This task does not produce a contractible signal. Thus, the agent is multi-tasking, but the principal observes only the outcome of one task before compensating the agent.[4]

The formal contract offered by the principal combines with the promise of external rewards to form a mixed stew of incentives that ultimately determines how hard the agent works on both tasks. In a recent survey of behavioral contract theory, Kőszegi (2014) singles out "the literature on the interaction between extrinsic and intrinsic motivation [as] one of the most exciting and productive in behavioral contract theory." In this literature, intrinsic motivation refers to non-monetary reasons why the agent would work hard on behalf of the principal. The call for more research is accompanied by the observation that "unlike extrinsic motivation, intrinsic motivation is a complex

[4]The term "action" should in the following be understood to refer to the *pair* of efforts devoted to the two different tasks. Conversely, a "task" describes one particular dimension of the action.

multifaceted phenomenon that is poorly understood." In the current paper, it is also the case that the contract does not capture all that is payoff-relevant to the agent. Thus, one "facet" of what looks like "intrinsic motivation" to the outsider may be that the agent has to evaluate how rewards on the job interact with private rewards. The level of risk the agent faces in the outside world, and his ability to influence it, may partly determine his effort on the job. Conversely, his employment contract may affect his drive to pursue private rewards. The agent faces a joint decision problem.

Indeed, the model can be interpreted as endogenizing the agent's pursuit of "work-life balance." The standard single-task model essentially focuses on the work dimension. In that model, the cost function may capture foregone leisure. However, given the separability that is usually assumed, the value of "leisure" is determined solely by effort at work but is otherwise independent of the contract terms. The current model, in contrast, allows the agent to simultaneously invest in both dimensions – "work" and "life" – while recognizing that the contract may be a deciding factor for both. A rewarding home life takes effort too. The implied multi-tasking turns out to alter some key predictions of standard contract theory.

The dominant method for analyzing moral hazard is the first-order approach (FOA). The FOA has been justified in a class of multi-tasking problems only very recently; see Kirkegaard (2014).[5] In this paper, I build on this work to extend the FOA to handle private rewards. Although rewards are assumed to be stochastically independent, the model allows for interdependencies in two ways. First, effort costs may be non-separable in the two tasks. Second, as alluded to before, rewards from the two different sources may be substitutes in the agent's utility function.

Thus, the first contribution of the paper is to present a tractable model of contracting in the presence of private rewards. The second contribution – justifying the FOA – is methodological in nature. That is, I provide a solution technique that can be used in future research on contracting with private rewards. Finally, the analysis provides new economic insights into at least three significant issues. First, I identify a new source of economic rents for the agent. Thus, unlike in the standard model, the agent may earn more than his reservation utility. Second, I show that contracts that may seem to have "flatter" incentives – yielding a smaller return to a marginal increase in on-the-job effort – may induce the agent to work *harder* on the job. Natu-

[5] Holmström and Milgrom (1988, 1991) present a specialized multi-tasking model in which the FOA is valid. This model relies on specific function forms and restricts contracts to be linear.

rally, this finding is at odds with the conventional wisdom in classical contract theory. Nevertheless, Kőszegi (2014) reviews several "behavioral economics" models in which results in this vein are obtained. Here, I identify a new mechanism, centered on considerations of "work-life balance", which is responsible for the result. Third, in the case of common agency, a new type of distortion that is absent in simpler models is uncovered. This distortion may in turn change the incentive for competing principals to share information among themselves.

In some situations, the non-contractible uncertainty may be payoff relevant to both the agent and the principal. Consider for example an entrepreneur (the principal) who has developed a novel product idea. However, the entrepreneur does not have the practical skills or engineering know-how to produce the product. Thus, he partners with an engineer (the agent), who is tasked with implementing the idea. The quality of the resulting product is verifiable, and can thus be contracted upon. However, both parties also care about status. The engineer may privately seek to bolster his status by claiming undue credit for the idea itself, thereby indirectly diminishing the entrepreneur's status. Even if the engineer is observed to be bragging, it is unlikely that such behavior is verifiable and contractible. Likewise, it is hard to quantify and contract upon status. It is in such cases, where the principal cares about both dimensions of the agent's action, that the optimal contract may deliver the agent more than his reservation utility.[6] As another example, consider the professional athlete whose unsavory conduct off the field impacts the team's reputation and thus its ticket and merchandise sales as well as its ability to attract talent in the future.[7]

In the standard moral hazard model, even with multi-tasking, it is well known that the agent's participation constraint must bind regardless of which action the principal seeks to implement (in the absence of a minimum wage). However, with private rewards, there are environments where some actions can only be implemented by awarding the agent with economic rents. In other words, for some actions the incentive compatibility constraints may restrict the set of feasible contracts so much that there is not enough remaining flexibility to make the participation constraint

[6]For technical reasons, I assume in the model (as in the example) that the principal's and agent's preferences over the private task are opposed. This is trivially satisfied when the principal does not directly care about the private task. Moreover, the assumption can be relaxed in special cases.

[7]The difference between the two examples is that in the latter it may be possible to contract upon ticket and merchandise sales. Note, however, that these signals depend on the player's actions both on and off the field. The model accomodates some such settings; see footnote 12, below.

binding as well; it is made redundant. It turns out that when the principal takes an interest in both tasks (but typically not otherwise), it may be optimal to implement precisely such an action.

It should be emphasized that this source of economic rents for the agent is new. Note that these rents are not due to the outside reward per se, but rather to the fact that this reward is private. After all, if the outside reward is contractible, the principal is effectively able to “take away” or appropriate the monetary value of the private reward by making the agent’s pay contingent on both signals.[8]

However, the multi-tasking aspect of the agent’s problem is directly responsible for another economically significant property. To illustrate, assume that the private reward is monetary as well, and that the agent’s preferences exhibits constant absolute risk aversion over total income. In this case, the agent’s expected utility (gross of effort costs) can be written in a convenient multiplicative form, as $M(a_1|\omega)N(a_2)$, where ω denotes the contract, a_1 effort on the job, and a_2 effort in pursuit of outside income. In the standard model, $N > 0$ is an exogenous constant. Consider now the set of contracts that induce the agent to deliver effort a_1 at work. Let ω' and ω'' denote two such contracts. In this model, it turns out that $M(a_1|\omega'') > M(a_1|\omega')$ if and only if $M'(a_1|\omega'') < M'(a_1'|\omega')$, where $M'(a_1|\omega)$ denotes the derivative of M with respect to a_1. To an outsider who fails to recognize that N is endogenous, it would thus seem that the ω'' delivers higher “base utility” than ω' but weaker explicit or extrinsic incentives. Nevertheless, the agent works just as hard with the latter contract. Thus, the outsider might be led to believe that one of the behavioral models reviewed by Kőszegi (2014) are at play.

For instance, Englmaier and Leider (2012) note that if the agent has reciprocal preferences, the principal can “generate intrinsic motivation” by giving the agent higher base utility. The agent then reciprocates by maintaining high effort even if explicit incentives are weakened. In the simplest version of Bénabou and Tirole’s (2003) model, the agent derives utility (a source of intrinsic motivation) if he performs well on the job. However, at the time of contracting the agent only has an imperfect

[8] There are other moral hazard models in which the agent earns more than his reservation utility. This is a typical property of a common model in which the agent is risk neutral but protected by limited liability. Moreover, Laffont and Martimort (2002, Section 5.3) explains how the agent may earn rents when his utility function is non-separable in income and effort. Note, however, that in this case the participation constraint is not redundant – it is just not optimal to make it binding. Moreover, in the current paper I follow most of the literature by assuming separability.

signal about the cost of effort. If the principal knows that effort is very costly, he may be worried that the agent has received a bad signal. Consequently, he is more likely to offer steeper explicit incentives to partially compensate, yet it may not be enough to prevent the probability of high effort from declining. Bénabou and Tirole (2003) note that if these considerations are not taken into account, the "outside observer might actually underestimate the power of these incentives [and] conclude that rewards are negative reinforcers."

In contrast, in the current model, the resolution to the puzzle is instead that the agent adjusts his effort at home when presented with ω'' instead of ω'. Specifically, the higher base utility at work leads the agent to work less hard at home. As labor income then plays a more significant role in the agent's overall well-being, weaker incentives are sufficient to induce him to work harder on the job. Similarly, let $\underline{\omega}'$ and $\underline{\omega}''$ denote the cheapest way of inducing a_1' and a_1'', respectively, with $a_1'' > a_1'$. I will show that if marginal costs are not too steep, then $M(a_1''|\underline{\omega}'') > M(a_1'|\underline{\omega}')$ whereas $M'(a_1''|\underline{\omega}'') < M'(a_1'|\underline{\omega}')$. While an outsider might explain the former by appealing to higher effort costs, the puzzle remains that the agent's utility appears to be less responsive to a marginal increase in effort under $\underline{\omega}''$. Again, the latter might be interpreted as indicating that explicit incentives are weaker when the agent works harder. As before, however, this ignores that the intensity with which the agent pursues private rewards has changed as well.

In the final part of the paper, the model is applied to common agency in which the agent works for two principals. The seminal paper on common agency with moral hazard is Bernheim and Whinston (1986), who assume that principals have access to the same information before rewarding the agent. Holmström and Milgrom (1987) consider a specialized setting – based on the linear-exponential-normal (LEN) model – which is sufficiently tractable that it is possible to relax the assumption that principals share the same information. In this case, each principal essentially views the agent as receiving private rewards in the form of some unknown monetary remuneration from the other principal. Thus, the theoretical contribution in the first part of the paper enables a reexamination of the common agency problem studied by Holmström and Milgrom (1987). Their model turns out to be not entirely robust. More accurately, I identify an additional distortion due to common agency which is absent in the simpler LEN model.[9] In short, the LEN model may underestimate the

[9] This observation echoes findings in Kirkegaard (2014), who documents that the LEN model's

aggregate cost of implementation under common agency. Due to this extra distortion, it may also be necessary to reevaluate the principals' incentives to share information, if possible. This issue is discussed briefly as well.

The remainder of the paper is organized as follows. Sections 2 and 3 analyze the problem from the agent's and principal's perspective, respectily, culminating in a justification of the FOA. Technical extensions are discussed in Section 4. Section 5 examines a particularly tractible environment. It is established that the agent may earn economic rents. The shape and properties of the optimal contract is discussed. Section 6 reviews some implications of the analysis, as well as the application to common agency. Section 7 concludes.

2 The agent's problem

Before describing the model and embarking on the analysis, I briefly outline some of the key steps on the road to justifying the FOA when allowing for private rewards. From a technical point of view, the multi-task justifications of the FOA presented by Kirkegaard (2014) are a good starting point for the current endeavour. In particular, one of his justifications apply to environments in which the contract is monotonic and such that the outcome of the two tasks endogenously appear as substitutes in the agent's utility function. That is, a marginal improvement in the performance of task 2 is worth less if the agent performed extremely well on task 1. The difference between Kirkegaard's (2014) model and the current model is that in the former it is the principal who rewards both tasks. However, in the examples in the beginning of this introduction, it is quite natural to assume that income earned by working for the principal is a substitute for private rewards. In other words, private rewards yields substitutability essentially for free. On the other hands, it turns out that private rewards make is substantially harder to establish that wages are monotonic.

To establish monotonicity, a main challenge is to sign the multipliers of the incentive compatibility constraints. This is accomplished by extending a classic argument by Rogerson (1985), involving a doubly-relaxed maximization problem. In essence, Rogerson (1985) shows that in the single-task model, it is sufficient to prevent the agent from working less hard than intended. In the present setting, however, there

predictions are not robust in settings involving multi-tasking. Note that the LEN model can handle monetary private rewards, whereas the current model permits non-monetary rewards as well.

are two tasks. As is perhaps intuitive, it turns out to be sufficient to simultaneously prevent the agent from shirking on job and working too hard on the private task.

The assumptions that are imposed on the primitives (technolocy and preferences) are thus used in various ways to either establish monotonicity and substitutability or to prove that the FOA is valid whenever the candidate contract takes such a form.

I consider the simplest possible model of a principal-agent relationship with private rewards that allows the agent to choose how hard to work on acquiring private rewards. The agent performs two "tasks", a_1 and a_2, each of which belong to a compact interval, $a_i \in [\underline{a}_i, \overline{a}_i]$, $i = 1, 2$. The first task captures the agent's effort on the job, as a result of which a contractible signal, x_1, is produced. The signal's marginal distribution is $G^1(x_1|a_1)$. The second "task" reflects the agent's pursuit of a private reward. The agent receives a (possibly non-monetary) reward, x_2, which is determined by the marginal distribution function $G^2(x_2|a_2)$. Here, a_2 could measure life-style choices and x_2 the health outcome. Assume x_i belongs to a compact interval, $[\underline{x}_i, \overline{x}_i]$, which is independent of a_i. Assume G^1 and G^2 are continuously differentiable in both variables to the requisite degree. Let $g^1(x_1|a_1)$ and $g^2(x_2|a_2)$ denote the respective densities. Assume $g^i(x_i|a_i) > 0$ for all $x_i \in [\underline{x}_i, \overline{x}_i]$ and all $a_i \in [\underline{a}_i, \overline{a}_i]$.[10] Note that each marginal distribution depends only on one task.[11] This property is further strengthened by assuming that x_1 and x_2 are independent.

Assumption A1 (Independence): *Outcomes are independent*, i.e. given a_1 and a_2, the joint distribution is given by

$$F(x_1, x_2|a_1, a_2) = G^1(x_1|a_1)G^2(x_2|a_2). \tag{1}$$

More structure is required on the components of the joint distribution function. Thus, define $l^i(x_i|a_i) = \ln g^i(x_i|a_i)$ and let $l^i_{a_i}(x_i|a_i)$ denote the likelihood-ratio, i.e. the derivative of $l^i(x_i|a_i)$ with respect to a_i, $i = 1, 2$.

Assumption A2 (MLRP): The marginal distributions have the *monotone likelihood*

[10] Throughout, all exogenous functions are assumed to be continously differentiable to the requisite degree. For brevity, statements to that effect are omitted from the numbered assumptions.

[11] This is somewhat less restrictive than it appears at first glance. For instance, assume G_i is a *one-parameter* distribution, and that a_1 and a_2 both influence the parameter. That is, G_i can be written $G_i(x_i|t_i(a_1, a_2))$. In this case, the problem can simply be reformulated to make t_1 and t_2 the two choice variables. However, the possibility that a_1 and a_2 influence different parameters of one or both of the marginal distributions is ruled out.

ratio property, i.e. for all $a_i \in [\underline{a}_i, \overline{a}_i]$ it holds that

$$\frac{\partial}{\partial x_i}\left(l^i_{a_i}(x_i|a_i)\right) = \frac{\partial^2 \ln g^i(x_i|a_i)}{\partial a_i \partial x_i} \geq 0$$

for all $x_i \in [\underline{x}_i, \overline{x}_i]$, with strict inequality on a subset of strictly positive measure, $i = 1, 2$.

Assumption A2 implies that $G^i_{a_i}(x_i|a_i) < 0$ for all $x_i \in (\underline{x}_i, \overline{x}_i)$.[12] The interpretation is that when the agent works harder, bad outcomes become less likely. In particular, if $a'_i > a''_i$ then $G^i(x_i|a'_i)$ first order stochastically dominates $G^i(x_i|a'_i)$.[13]

It is assumed that x_1 and x_2 are realized at the same time. In an important paper, Rogerson (1985) justifies the FOA in a one-signal, one-task model. He assumes the distribution function satisfies MLRP and that it is convex in the (one-dimensional) action. Rogerson (1985) refers to the latter as the convexity of distribution function condition (CDFC). Kirkegaard (2014) extends the justification of the FOA to allow multiple tasks and signals. In the process, he shows that a natural extension of the CDFC is to assume that the distribution function is convex in the (now many-dimensional) action. The same assumption is imposed here.

Assumption A3 (LOCC): $F(x_1, x_2|a_1, a_2)$ satisfies the *lower orthant convexity condition*, i.e. $F(x_1, x_2|a_1, a_2)$ is weakly convex in (a_1, a_2) for all (x_1, x_2) and all (a_1, a_2).

Assumption A3 necessitates that G^i is convex in a_i, $i = 1, 2$. In fact, it implies that $G^i_{a_i a_i}(x_i|a_i) > 0$ for all $x_i \in (\underline{x}_i, \overline{x}_i)$.[14] A sufficient condition for LOCC is that G^1 and G^2 are both log-convex. Kirkegaard (2014) lists several examples of log-convex distribution functions. See also Ábrahám et al (2011), discussed in Section 6.4. Alternatively, fix some G^1 that is strictly convex in a_1, but not necessarily log-convex. Then, there is always some "sufficiently convex" G^2 function that ensures that Assumption A3 is satisfied. For example, a non-negative function $h(z)$ is said to be ρ-convex if $h(z)^\rho/\rho$ is convex, or $h''(z)h(z)/h'(z)^2 \geq 1 - \rho$ for all z. Thus,

[12] To see this, recall first that the expected value of $l^i_{a_i}(x_i|a_i)$ is zero. Assumption A2 therefore implies that $l^i_{a_i}(\underline{x}_i|a_i) < 0 < l^i_{a_i}(\overline{x}_i|a_i)$. Since $G^i_{a_i}(\underline{x}_i|a_i) = G^i_{a_i}(\overline{x}_i|a_i) = 0$, it follows that $G^i_{a_i}(x_i|a_i) = \int_{\underline{x}_i}^{x_i} l^i_{a_i}(z_i|a_i) g^i(x_i|a_i) < 0$ for all $x_i \in (\underline{x}_i, \overline{x}_i)$.

[13] The model would reduce to the standard single-task, one-signal model if $G^2(x_2|a_2)$ was degenerate and independent of a_2.

[14] LOCC necessitates that $G^i_{a_i a_i} \geq 0$ and $G^1 G^2 G^1_{a_1 a_1} G^2_{a_2 a_2} - \left(G^1_{a_1} G^2_{a_2}\right)^2 \geq 0$. At any interior (x_1, x_2), the last term is strictly positive, by A2. Thus, it is necessary that $G^1_{a_1 a_1} > 0$ and $G^2_{a_2 a_2} > 0$.

a ρ-convex function is log-convex if and only if $\rho \leq 0$ (and convex if and only if $\rho \leq 1$). It is easy to see that if $G^2(x_2|a_2)$ satisfies Assumption A2 and is ρ-convex in a_2 (for all x_2) for some small enough ρ (i.e. ρ is negative, but numerically large), then Assumption A3 is satisfied.[15] The point is that as long as G^1 satisfies a strict version of CDFC, then there are G^2 functions that will permit the FOA to be justified even when allowing for private rewards.

Assumptions A1–A3 describes the "technology". The next set of assumptions describes the agent's preferences. Given action (a_1, a_2), wage w, and private reward x_2, the agent's utility is assumed to take the form

$$v(w, x_2) - c(a_1, a_2),$$

where v is a benefit function and c a cost function. Both functions are assumed to be continuously differentiable in both their arguments to the requisite degree. The function $v(w, x_2)$ is strictly increasing and strictly concave in both arguments, $v_i > 0 > v_{ii}$, $i = 1, 2$, where subscripts denote derivatives. The cost function is likewise assumed to be strictly increasing. It is also assumed to be convex in (a_1, a_2). While Assumption A1 imply that there is no stochastic interaction between a_1 and a_2, the cost function allows interaction between tasks.

Note that if the private reward, x_2, is income, then $v(w, x_2)$ could be written $v(w+x_2)$, in which case it is automatic that $v_{12} = v_{11} = v_{22} < 0$. That is, employment income and outside income are substitutes. Indeed, even when x_2 is not income it is natural to assume that w and x_2 are strict substitutes. Thus, it will be assumed that $v_{12} < 0$; the higher x_2 is, the lower is the marginal utility of additional employment income. I will also assume that a_1 and a_2 are weak substitutes in the cost function, or $c_{12} \geq 0$. That is, the marginal cost of increasing a_1 is higher the higher a_2 is. Assumption A4 summarizes these assumptions on the agent's preferences.

Assumption A4 (Substitutes): The agent's Bernoulli utility is $v(w, x_2) - c(a_1, a_2)$; $v(w, x_2)$ is strictly increasing and strictly concave in both w and x_2, while $c(a_1, a_2)$ is strictly increasing and weakly convex in (a_1, a_2). The rewards w and x_2 are strict substitutes; $v_{12}(w, x_2) < 0$. The tasks are weak substitutes; $c_{12}(a_1, a_2) \geq 0$.

[15] The inequality in the previous footnote can be written $G^1 G^1_{a_1 a_1} (G^2 G^2_{a_2 a_2} / \left(G^2_{a_2}\right)^2) - \left(G^1_{a_1}\right)^2 \geq 0$, for interior (x_1, x_2). By ρ-convexity, the left hand side is greater than $G^1 G^1_{a_1 a_1} (1 - \rho) - \left(G^1_{a_1}\right)^2 \geq 0$. Hence, the inequality is satisfied if ρ is small enough.

The principal specifies a contract of the form $w(x_1)$. That is, the contract details the wage to the agent if the verifiable signal is x_1. Upon taking action (a_1, a_2), the agent's expected payoff is then

$$EU(a_1,a_2) = \int\int v(w(x_1),x_2)g^1(x_1|a_1)g^2(x_2|a_2)dx_1dx_2 - c(a_1,a_2).$$

Imagine the principal is designing the incentive schme $w(x_1)$ with the intention of inducing the agent of take action (a_1', a_2'). For the agent to comply, his expected utility must be maximized at (a_1', a_2'). Assuming the action is interior, this at the very least necessitates that expected utility is at a stationary point at (a_1', a_2'), or $EU_1(a_1', a_2') = EU_2(a_1', a_2') = 0$. The FOA relies on the latter conditions being not only necessary but also sufficient for utility maximization. If this is the case, one need not worry about the agent deviating from (a_1', a_2'). To this end, the typical approach is to establish that $EU(a_1, a_2)$ is concave, although Kirkegaard (2014) uses a somewhat different method.

Note that if $w(x_1)$ is non-decreasing in x_1, then, given Assumption A4, $v(w(x_1), x_2)$ is increasing in both x_1 and x_2, and submodular in the two. This is noteworthy, because Kirkegaard (2014) proves that if the agent faces such a reward function, then the FOA is valid if LOCC (Assumption A3) is satisfied as well.[16] In fact, $EU(a_1, a_2)$ is concave in (a_1, a_2). To see this, note that after integration by parts with respect to x_2,

$$EU(a_1,a_2) = \int \left(v(w(x_1),\overline{x}_2) - \int v_2(w(x_1),x_2)G^2(x_2|a_2)dx_2 \right) g^1(x_1|a_1)dx_1 - c(a_1,a_2) \tag{2}$$

Assuming for simplicity that $w(x_1)$ is differentiable (it will later be established that the optimal contract is indeed differentiable), another round of integrating by parts,

[16] Kirkegaard (2014) uses insights from choice under uncertainty to explain the intuition underlying this result. Jewitt (1988) and Conlon (2009) present two results with a similar flavour in a model with two signals but a single task. Kirkegaard's (2014) characterization is more general, as it extends to more signals and more tasks. For instance, he identifies the appropriate generalization of submodularity when there are more than two signals.

this time with respect to x_1, yields

$$\begin{aligned} EU(a_1, a_2) &= v(w(\overline{x}_1), \overline{x}_2) - \int v_1(w(x_1), \overline{x}_2) w'(x_1) G^1(x_1|a_1) dx_1 \\ &\quad - \int v_2(w(\overline{x}_1), x_2) G^2(x_2|a_2) dx_2 \\ &\quad + \int \int v_{12}(w(x_1), x_2) w'(x_1) G^1(x_1|a_1) G^2(x_2|a_2) dx_1 dx_2 - c(a_1, a_2) \end{aligned} \quad (3)$$

Recall that $v_1, v_2 > 0 \geq v_{12}$ while $G^1(x_1|a_1)$, $G^2(x_2|a_2)$, $G^1(x_1|a_1)G^2(y_2|a_2)$, and $c(a_1, a_2)$ are all convex in $(a_{1,}a_2)$. Thus, as long as $w'(x_1) \geq 0$, $EU(a_1, a_2)$ is the sum of concave functions. The first Lemma records this fact

Lemma 1 *Assume $w'(x_1) \geq 0$ for all $x_1 \in [\underline{x}_1, \overline{x}_1]$ and that Assumptions A1–A4 hold. Then, the agent's expected utility, $EU(a_1, a_2)$, is concave in (a_1, a_2).*

Unfortunately, it is far from trivial to establish that $w(x_1)$ is non-decreasing. Thus, most of the following analysis is devoted to that particular problem. The previous assumptions will also play a role in establishing that property, but further (economically meaningful) assumptions are also needed.

First, however, note that Lemma 1 implies that $EU(a_1, a_2)$ must be concave in a_1 and concave in a_2, whenever $w'(x_1) \geq 0$. In fact, (2) implies that $EU_{22}(a_1, a_2) < 0$ as $G^2_{a_2 a_2} > 0$. The expression in (3) also makes it clear that the last two parts of Assumption A4 pull in the same direction. In particular, the two tasks are strict substitutes in the agent's expected utility, or $EU_{12}(a_1, a_2) \leq 0$, whenever $w'(x_1) \geq 0$. Indeed, if the inequality is strict on a subset of positive measure, then $v_{12} < 0$ and $c_{12} \geq 0$ imply that $EU_{12}(a_1, a_2) < 0$.

To proceed, it is necessary to impose more specific assumptions on the agent's risk preferences over labor income, w. Thus, it will be assumed that the absolute risk aversion with respect to w is decreasing in x_2. In other words, the agent is less sensitive to risk in labor income the higher the private reward is.

Assumption A5 (Decreasing absolute risk aversion): The agent's absolute risk aversion over labor income is decreasing in x_2. That is, $v_1(w, x_2)$ is log-supermodular in (w, x_2), or

$$\frac{\partial^2 \ln v_1(w, x_2)}{\partial w \partial x_2} \geq 0 \text{ for all } w \text{ and all } x_2 \in [\underline{x}_2, \overline{x}_2]. \quad (4)$$

Of course, (4) is equivalent to

$$\frac{\partial}{\partial x_2}\left(\frac{-v_{11}(w,x_2)}{v_1(w,x_2)}\right) \leq 0.$$

For future reference, note that Assumption A2 (MLRP) is equivalent to the requirement that $g^i(x_i|a_i)$ is log-supermodular in (x_i, a_i), $i = 1, 2$. To continue, let

$$V(w, a_2) = \int v(w, x_2) g^2(x_2|a_2) dx_2, \tag{5}$$

such that the agent's expected utility can be written as

$$EU(a_1, a_2) = \int V(w(x_1), a_2) g^1(x_1|a_1) dx_1 - c(a_1, a_2).$$

Given (5), note that

$$\begin{aligned} V_1(w, a_2) &= \int v_1(w, x_2) g^2(x_2|a_2) dx_2 > 0, \text{ and} \\ V_{12}(w, a_2) &= \int v_1(w, x_2) g^2_{a_2}(x_2|a_2) dx_2 < 0. \end{aligned}$$

Here, $V_1(w, a_2)$ describes the expected marginal utility of additional labor income given the agent's effort on the private task is a_2. Of course, $V_{12}(w, a_2)$ captures how this expectation changes with a_2. Assumptions A1 and A4 together implies that $V_{12}(w, a_2) < 0$. Evidently, $V_1(w, a_2)$ is strictly decreasing in w, or $V_{11}(w, a_2) < 0$. Moreover, the term under the integration sign in $V_1(w, a_2)$ is, by Assumptions A2 and A5, log-supermodular in (w, x_2, a_2). As described by e.g. Athey (2002), log-supermodularity is preserved under integration. Thus, $V_1(w, a_2)$ is log-supermodular in (w, a_2). That is, the agent's decreasing absolute risk aversion aggregates, or carries over to the expected utility in (5). In other words,

$$\frac{\partial}{\partial a_2}\left(\frac{-V_{11}(w,a_2)}{V_1(w,a_2)}\right) \leq 0,$$

such that the agent is less sensitive to risk in labor income the better the distribution of private rewards is. Equivalently,

$$\frac{\partial}{\partial w}\left(\frac{-V_{12}(w,a_2)}{V_1(w,a_2)}\right) \leq 0. \tag{6}$$

The latter property is especially important. Technically, the assumption that $g^2(x_2|a_2)$ is log–supermodular can be relaxed (although $G^2_{a_2} \leq 0$ is still required) if the assumption that (4) holds is replaced by the assumption that (6) holds. For instance, if x_2 is income and $v(w, x_2) = -e^{-r(w+x_2)}$, $r > 0$, then the agent exhibits constant absolute risk aversion in total income (and its components). In this case, (6) is trivially satisfied for any $g^2(x_2|a_2)$.

To complete the description of the agent's problem, assume that the only constraint (other than incentive compatibility) is an individual rationality (participation) constraint. That is, the agent must be assured expected utility of at least $\overline{u}$ to sign the contract.

3 Contracts with private rewards

The previous section focused on describing the problem from the agent's point of view. Consider now the principal's problem. First, assume that the principal is risk neutral and that his utility depends directly on a_1 and, possibly, on a_2. Let $B(a_1, a_2)$ denote this direct benefit, and assume that it is continuously differentiable. For instance, $B(a_1, a_2)$ could be the expected value of x_1, given a_1. As explained below, for technical reasons I assume that $B_2(a_1, a_2) \leq 0$, such that the principal prefers a_2 to be as small as possible. This assumption is of course satisfied if B is independent of a_2. In the more specialized model in Section 4, it is possible to allow $B_2(a_1, a_2) > 0$, however. Finally, let $E[w|a_1, a_2]$ denote the expected wage costs if the agent is induced to take action (a_1, a_2).

Assumption A6 (The principal's preferences): The principal is risk neutral, with expected utility $B(a_1, a_2) - E[w|a_1, a_2]$, where $B_2(a_1, a_2) \leq 0$ for all (a_1, a_2).

It is natural to assume that $B(a_1, a_2)$ is increasing in a_1. Indeed, Proposition 2 in Section 6.2 will establish that if it is optimal to implement an interior action, then $B_1 > 0$ at that point. However, it is not necessary that B be globally increasing in a_1.

The principal's problem is to maximize $B(a_1, a_2)$ less wage costs, subject to individual rationality and incentive compatibility, or

$$\max_{a_1,a_2,w} B(a_1,a_2) - \int w(x_1)g_1(x_1|a_1)dx_1$$
$$st. \quad EU(a_1,a_2) \geq \overline{u}$$
$$(a_1,a_2) \in \arg \max_{(a_1',a_2')\in[\underline{a}_1,\overline{a}_1]\times[\underline{a}_2,\overline{a}_2]} EU(a_1',a_2')$$

The action that solves the problem is henceforth referred to as the second-best action.

It is important to realize that the contract indirectly determines not only how hard the agent works for the principal, but also how hard he works on the private task. From the agent's point of view, the function $v(w(x_1),x_2)$ is crucial to the decision of how much effort to devote to each task. In fact, for incentive compatibility, it is immaterial that the reward x_2 happens to be not paid by the principal.

Assume the principal wishes to induce an interior action. Then, as mentioned previously, it is necessary that EU achieves a stationary point at the targeted action, or $EU_1(a_1,a_2) = 0 = EU_2(a_1,a_2)$. These constraints are referred to as the "local" incentive compatibility constraints. Of course, the local constraints do not rule out that the agent finds it optimal to deviate to actions further away. As in the existing FOA literature, the main objective of this part of the paper is to establish conditions under which the local constraints are in fact sufficient for "global" incentive compatibility.

Consider the following *relaxed problem*, so named because the incentive compatibility constraint in the original problem has been relaxed,

$$\max_{a_1,a_2,w} B(a_1,a_2) - \int w(x_1)g_1(x_1|a_1)dx_1$$
$$st. \quad EU(a_1,a_2) \geq \overline{u}$$
$$EU_1(a_1,a_2) = 0$$
$$EU_2(a_1,a_2) = 0$$

The FOA is said to be valid if the solution to the relaxed problem also solves the unrelaxed (original) problem.

The description of the problem thus far does not rule out that wages are at a corner of the domain of $v(w,x_2)$. For simplicity, it will be assumed that wages are interior. Rogerson (1985) impose assumptions directly on the utility functions (see his assumption A3–A4 and A6–A7) to achieve this.

Assumption A7 (Interior wages): Any solution to the relaxed problem involves only wages in the interior of the domain of $v(w, x_2)$.

Let $\lambda \geq 0$ denote the multiplier to the participation constraint, and μ_1 and μ_2 denote the multipliers to the two local incentive compatibility constraints in the relaxed problem. Given Assumption A7, the optimal wage if x_1 is observed is implicitly characterized by the necessary first order condition,

$$V_1(w, a_2)\left[\lambda + \mu_1 l^1_{a_1}(x_1|a_1)\right] + \mu_2 V_{12}(w, a_2) = 1 \tag{7}$$

or

$$\lambda + \mu_1 l^1_{a_1}(x_1|a_1) = \frac{1}{V_1(w, a_2)} - \mu_2 \frac{V_{12}(w, a_2)}{V_1(w, a_2)}. \tag{8}$$

Qualitatively, the solution almost certainly depends on the sign of the two multipliers μ_1 and μ_2. Indeed, it is not even clear that there is a unique solution. However, later arguments will establish that it is sufficient to focus on multipliers for which $\mu_1 \geq 0 \geq \mu_2$. Then, the contract $w(x_1)$ is well-behaved.

Lemma 2 *Assume that Assumptions A1-A7 hold. Then, $w(x_1)$ is unique in any solution to the relaxed problem for which $\mu_1 \geq 0 \geq \mu_2$. Moreover, in this case the solution is differentiable, with $w'(x_1) \geq 0$ for all $x_1 \in [\underline{x}_1, \overline{x}_1]$. If $\mu_1 > 0 \geq \mu_2$, then $w'(x_1) > 0$ for a subset of $[\underline{x}_1, \overline{x}_1]$ of strictly positive measure.*

Proof. Given $\mu_2 \leq 0$, $V_{11} < 0$ and (6) imply that the right hand side of (8) is strictly increasing in w. Thus, for each x_1 there is at most one solution to (8), $w(x_1)$. Differentiablity then follows automatically from the differentiability of all the components in (8). Since $\mu_1 \geq 0$, Assumption A2 (MLRP) implies that the left hand side is non-decreasing in x_1. Hence, $w(x_1)$ is non-decreasing in x_1. The last part likewise follows from Assumption A2. ∎

The next step is to show that the optimal multipliers indeed satisfy $\mu_1 \geq 0 \geq \mu_2$. The proof utilizes Rogerson's idea of considering a *doubly-relaxed* problem. In Rogerson's one-task model, the relaxed incentive compatibility constraint, $EU_1 = 0$, is replaced with the even weaker constraint that $EU_1 \geq 0$. In the current multi-task model, the appropriate doubly-relaxed problem weakens the relaxed problem studied above by replacing $EU_1 = 0$ and $EU_2 = 0$ with $EU_1 \geq 0$ and $EU_2 \leq 0$, respectively. Note that in the doubly-relaxed problem, it is thus the case that $\mu_1 \geq 0 \geq \mu_2$.

As in Rogerson, assume there is a solution to the doubly-relaxed problem (which is necessary for the existence of a solution to the relaxed problem). As before, assume any solution is interior.

Assumption A8 (The doubly-relaxed problem): A solution to the doubly-relaxed problem exists. Any solution involves only wages in the interior of the domain of $v(w, x_2)$.

Of course, any solution to the doubly-relaxed problem must take the form in (8). By Lemma 2, any solution thus features non-decreasing wages. The crucial step is to prove that the solution to the doubly-relaxed problem coincides with the solution to the relaxed problem. To begin, this necessitates that the constraints $EU_1(a_1, a_2) \geq 0$ and $EU_2(a_1, a_2) \leq 0$ are binding, which is certainly the case if $\mu_1 > 0 > \mu_2$.

Lemma 3 *Assume the solution to the doubly-relaxed problem involves an interior action. Then, given Assumptions A1–A8, the solution to the doubly-relaxed problem satisfies $\mu_1 > 0 > \mu_2$. Thus, the solution to the doubly-relaxed problem also solves the relaxed problem.*

Proof. The possibility that $\mu_1 = 0$ is easily ruled out, since it would imply a constant wage and thus $EU_1 < 0$, which violates the doubly-relaxed constraints. In other words, $\mu_1 > 0$ and so $EU_1(a_1, a_2) = 0$. Now, note that the adjoint equation for the optimal a_2 is

$$[B_2(a_1, a_2) + \lambda EU_2(a_1^*, a_2^*) + \mu_1 EU_{12}(a_1^*, a_2^*)] + \mu_2 EU_{22}(a_1^*, a_2^*) = 0. \tag{9}$$

By Assumption A6, $B_2(a_1, a_2) \leq 0$. Following the argument described after Lemma 1, it holds that $EU_{12}(a_1^*, a_2^*) < 0$ given the properties of $w(x_1)$ described in Lemma 2 when $\mu_1 > 0 \geq \mu_2$.[17] Since $\lambda EU_2(a_1^*, a_2^*) \leq 0$, the term in the bracket in (9) is thus strictly negative. As $EU_{22}(a_1^*, a_2^*) < 0$, it is therefore necessary that $\mu_2 < 0$. Hence, $EU_2(a_1^*, a_2^*) = 0$. Thus, both incentive constraints are binding. In other words, the solution is also feasible in the relaxed problem. Consequently, the solution to the double-relaxed problem also solves the relaxed problem. ■

[17]Note that Assumption A4 rules out that $v_{12} = c_{12} = 0$. However, this case seems relatively uninteresting, as it would imply that there is a fixed a_2 which is optimal for the agent regardless of the contract. In particular, the last term in (7) would disappear. It is then easy to show that the FOA is valid if $G^1(x_1|a_1)$ satisfies CDFC.

To finish, it is also necessary that there is no solution to the relaxed problem that is not a solution to the doubly-relaxed problem. However, any solution to the relaxed problem is also feasible in the doubly-relaxed problem. For this reason, the solutions to the two problems coincide.

Lemma 4 *Assume the solution to the doubly-relaxed problem involves an interior action. Then, given Assumptions A1–A8, the solutions to the relaxed and doubly-relaxed problems coincide. Hence,* $\mu_1 > 0 > \mu_2$ *in the relaxed problem.*

Proof. Lemma 3 proves that any solution to the doubly-relaxed problem also solves the relaxed problem. For the other direction, assume there is some solution to the relaxed problem that does not solve the doubly-relaxed problem. Evidently, this solution is feasible in both problems. Since it is not optimal in the doubly-relaxed problem, is must be dominated by the solution to the doubly-relaxed problem. However, by Lemma 3 the latter is also feasible in the relaxed problem, thereby contradiction the assumption that the original "solution" actually solves the relaxed problem. ■

Lemmata 2 and 4 imply that $w(x_1)$ is non-decreasing in the relaxed problem. Lemma 1 then implies that the FOA is valid as long as the second-best is in the interior.

Theorem 1 *Assume the second-best action* (a_1, a_2) *is interior. Then, given Assumptions A1–A8, the FOA is valid.*

It is natural to be sceptical about the assumption in the theorem that the principal finds it optimal to induce the agent to work on the private task (a_2 is interior). However, he may have no choice, especially if the agent's marginal costs are sufficiently low. More formally, note that

$$EU_2(a_1, \underline{a}_2) = \int \left(\int v(w(x_1), x_2) g^2_{a_2}(x_2|\underline{a}_2) dx_2 \right) g^1(x_1|a_1) dx_1 - c_2(a_1, \underline{a}_2).$$

The inner integral is strictly positive regardless of the contract. Thus, if $c_2(a_1, \underline{a}_2) = 0$, then $EU_2(a_1, \underline{a}_2) > 0$ regardless of the contract. In this case, it is impossible to persuade the agent to refrain from pursuing private rewards. Obviously, this is yet another indication that private rewards are relevant for the moral hazard problem, and worthy of study. Moreover, Section 5 considers a particularly tractable version of the model in which it can be shown that wage costs are in fact decreasing in a_2 up to a certain point.

Finally, the reason Lemma 1 holds is that the agent's utility is concave in (a_1, a_2) whenever $w(x_1)$ is non-decreasing.[18] Consequently, it is technically possible to handle boundary actions as well. For instance, if $EU_2(a_1, \underline{a}_2) \leq 0 = EU_1(a_1, \underline{a}_2)$, then the agent has no incentive to deviate from $(a_1, \underline{a}_2)$.

4 Extension: Relaxing LOCC

The stringency of Rogerson's CDFC is the main source of criticism of the FOA in the standard model. As mentioned earlier, Assumption A3 (LOCC) generalizes CDFC to allow multi-tasking. In this section, I consider one possible way to relax Assumption A3. The next section examines an even further relaxation of Assumption A3, which comes at the cost of imposing much more structure on $v(w, x_2)$. Indeed, this structure will also make it possible to derive a fairly complete description of the optimal (a_1, a_2) pair.

Jewitt (1988) was first to relax the CDFC. In the one-signal, one-task case, he replaces the CDFC assumption that $G^1(x_1|a_1)$ is convex in a_1 with the assumption that

$$\int_{\underline{x}_1}^{x_1} G^1(y_1|a_1)dy_1 \tag{10}$$

is convex in a_1 for all x_1. Jewitt (1988) shows that this condition is sufficient to justify the FOA provided that the agent's utility is increasing and concave in x_1. Concavity must thus be established. In Jewitt's setting, this turns out to require that the likelihood ratio is increasing and concave. Kirkegaard (2014) proves that this type of justification of the FOA can be extended to many signals and many tasks, as long as the tasks are independent. In the case with two signals and two tasks, the appropriate assumption is that the antiderivative of $F(x_1, x_2|a_1, a_2)$ is weakly convex in (a_1, a_2), as described in the next assumption. Kirkegaard (2014) terms this condition the *cumulative lower orthant convexity condition* (CLOCC).

ASSUMPTION A3' (CLOCC): $F(x_1, x_2|a_1, a_2)$ satisfies the *cumulative lower orthant*

[18] This assertion can be established more formally using similar steps as in the proof of Theorem 2, below.

convexity condition (CLOCC), i.e.

$$\int_{\underline{x}_1}^{x_1}\int_{\underline{x}_2}^{x_2} G^1(y_1|a_1)G^2(y_2|a_2)dy_2dy_1 = \int_{\underline{x}_1}^{x_1} G^1(y_1|a_1)dy_1 \int_{\underline{x}_2}^{x_2} G^2(y_2|a_2)dy_2 \tag{11}$$

is weakly convex in (a_1, a_2) for all (x_1, x_2) and all (a_1, a_2).

Unfortunately, with private rewards it seems infeasible to sign the second derivative of $v(w(x_1), x_2)$ with respect to x_1 in general. The next section considers a special case where it is possible, and where CLOCC may serve to justify the FOA.

However, it is straightforward to sign the second derivative of $v(w(x_1), x_2)$ with respect to x_2. In fact, by assumption, the agent's utility is concave in x_2. This observation opens the door for a simpler relaxation of Assumption A3. In particular, I will exploit that $v(w(x_1), x_2)$ is increasing in x_1 and increasing and concave in x_2. Thus, a hybrid of LOCC and CLOCC is called for.

Assumption A3'' (HOCC): $F(x_1, x_2|a_1, a_2)$ satisfies the *hybrid orthant convexity condition (HOCC)*, i.e.

$$G^1(x_1|a_1)\int_{\underline{x}_2}^{x_2} G^2(y_2|a_2)dy_2 \tag{12}$$

is weakly convex in (a_1, a_2) for all (x_1, x_2) and all (a_1, a_2).

Assumptions A3, A3', and A3" can be ordered according to how restrictive they are. Specifically, Assumption A3 (LOCC) implies A3" (HOCC), which in turn implies A3' (CLOCC). As was the case for Assumption A3, Assumptions A3' and A3" both imply that each term in (11) and (12), respectively, must be strictly convex in a_i for interior x_i. As before, HOCC is satisfied if e.g. the two terms are log-convex in a_1 and a_2, respectively. Although HOCC is weaker than LOCC, it is easy to see that it remains the case that $EU(a_1, a_2)$ is strictly concave in a_2, or $EU_{22}(a_1, a_2) < 0$, regardless of the contract. Formally, this can be established by using integration by parts twice and invoking the assumption that $v_{22} < 0$. Likewise, as before, $EU(a_1, a_2)$ is concave in a_1 whenever the contract is monotonic. While the agent's expected utility is thus concave in each task, it remains to show that it is jointly concave in (a_1, a_2).

As explained in Kirkegaard (2014), multi-signal justifications of the FOA that rely on LOCC or CLOCC also require one to sign certain cross-partial derivatives. For instance, recall that $v_{12} < 0$ was invoked to prove Lemma 1. Given HOCC, it turns out to be sufficient to add the mild assumption that $v_{122} \geq 0$. In other words, the price of weakening Assumption A3 by replacing it with Assumption A3" is that $v_{122} \geq 0$ must be assumed. However, note that if x_2 is income, then $v_{122} \geq 0$ is implied by Assumption A5. In this case, then, there is no cost of replacing A3 by A3". More generally, in view of Assumption 5, $v_{122} \geq 0$ is satisfied if $v_2(w, x_2)$ is log-supermodular in (w, x_2), i.e. if the agent's risk aversion with respect to the private rewards is decreasing with labor income.

Assumption A9 (Supermodular marginal utility): The agent's marginal utility of the private reward is supermodular. That is, $v_2(w, x_2)$ is supermodular in (w, x_2), or $v_{122}(w, x_2) \geq 0$.

Theorem 2 proves that the FOA remains valid once LOCC is replaced by HOCC, provided that Assumption A9 is imposed as well.

Theorem 2 *Assume the second-best action (a_1, a_2) is interior. Then, given Assumptions A1,A2,A3", and A4–A9, the FOA is valid.*

Proof. The argument that $w(x_1)$ is increasing in x_1 remains unchanged as Assumption A3 is replaced by Assumption A3". Using integration by parts repeatedly then yields

$$\begin{aligned} EU(a_1, a_2) &= v(w(\overline{x}_1), \overline{x}_2) - v_2(w(\overline{x}_1), \overline{x}_2) \int G^2(x_2|a_2)dx_2 \\ &+ \int v_{22}(w(\overline{x}_1), x_2) \int_{\underline{x}_2}^{x_2} G^2(y_2|a_2)dy_2 dx_2 \\ &- \int v_1(w(x_1), \overline{x}_2) w'(x_1) G^1(x_1|a_1)dx_1 \\ &+ \int v_{12}(w(x_1), \overline{x}_2) w'(x_1) \left(\int G^2(x_2|a_2)dx_2 G^1(x_1|a_1) \right) dx_1 \\ &- \int \int v_{122}(w(x_1), x_2) w'(x_1) \left(\int_{\underline{x}_2}^{x_2} G^2(y_2|a_2)dy_2 G^1(x_1|a_1) \right) dx_2 dx_1 \\ &- c(a_1, a_2). \end{aligned}$$

By assumption, $v_1, v_2 > 0 > v_{12}, v_{22}$ and $v_{112} \geq 0$. Moreover, by Assumption A3", every term involving a distribution function is convex in (a_1, a_2). Thus, $EU(a_1, a_2)$ is the sum of functions that are concave in (a_1, a_2). Hence, the agent's utility is concave.

Consequently, the agent has no incentive to deviate from the (a_1, a_2) pair implied by the "local" incentive compatibility constraints. ■

5 Multiplicative rewards

In this section, I consider a particularly tractable type of utility function. From a technical perspective, this makes it possible to further relax Assumption A3. In economics terms, it is significant that the added structure makes it possible to gain further insights into the optimal (a_1, a_2) pair. As a consequence of these results, I will demonstrate that the agent may earn more than his reservation utility in a principal-agent relationship with private rewards.

5.1 The first-order approach with multiplicative rewards

As mentioned above, imposing more structure on the reward function $v(w, x_2)$ makes it possible to replace Assumption A3 or A3" with the even weaker Assumption A3'. Thus, assume now that the reward function is multiplicative, in the sense that $v(w, x_2)$ can be written as

$$v(w, x_2) = -m(w)n(x_2), \tag{13}$$

where m and n are strictly negative functions that are strictly increasing and strictly concave on their domain. Note that Assumptions A5 and A9 are trivially satisfied, as is the part of Assumption A4 that pertains to $v(w, x_2)$. An obvious example that satisfies (13) is $m(w) = -e^{-rw}$ and $n(x_2) = -e^{-rx_2}$, for any $r > 0$. Then, $v(w, x_2) = -e^{-r(w+x_2)}$. Here, x_2 can be interpreted as income, and the agent exhibits constant absolute risk aversion.

For a given $w(x_1)$ contract, let

$$M(a_1) = \int m(w(x_1))g^1(x_1|a_1)dx_1 < 0 \tag{14}$$

$$N(a_2) = -\int n(x_2)g^2(x_2|a_2)dx_2 > 0, \tag{15}$$

such that

$$EU(a_1, a_2) = M(a_1)N(a_2) - c(a_1, a_2).$$

In this multiplicative case, (8) can be written in a much simpler form,

$$(\lambda N(a_2) + \mu_2 N'(a_2)) + (\mu_1 N(a_2))\, l^1_{a_1}(x_1|a_1) = \frac{1}{m'(w)}, \tag{16}$$

or, by renaming the terms in the parentheses,

$$\widehat{\lambda} + \widehat{\mu} l^1_{a_1}(x_1|a_1) = \frac{1}{m'(w)}$$

precisely as in the usual model with no private rewards. Standard methods can now be used to prove that $\mu_1 > 0$ (or $\widehat{\mu} > 0$). Specifically, given MLRP, the contract would be non-increasing if $\mu_1 \leq 0$, thus implying that $EU_1 < 0$ in violation of the incentive-compatibility constraints. In other words, given only that $a_1 > \underline{a}_1$, it must hold that $\mu_1 > 0$ such that the optimal contract is non-decreasing in x_1 regardless of which (a_1, a_2) the principal seeks to implement. Note that as long as wages are interior, this argument applies to all (a_1, a_2) pairs with $a_1 > \underline{a}_1$, not only the pair that turns out to be optimal. Note also that it is not necessary to sign μ_2. Thus, there is no need to consider the doubly-relaxed problem, and thus no need to invoke Lemma 3, which is the only place where the assumption that $B_2(a_1, a_2) \leq 0$ is utilized. In other words, in the multiplicative model, this part of Assumption A6 can be relaxed.

Assumption A6' (The principal's preferences): The principal is risk neutral, with expected utility $B(a_1, a_2) - E[w|a_1, a_2]$.

Given the similarities to the standard model, Jewitt's (1988) proof that $m(w(x_1))$ is sometimes concave in x_1 can be adopted without change. In particular, Jewitt proves that $m(w(x_1))$ is increasing and concave in x_1 if $l^1_{a_1}(x_1|a_1)$ is increasing and concave and

$$\frac{d}{dw}\left(\frac{-m''(w)}{m'(w)^3}\right) \geq 0. \tag{17}$$

The latter condition is satisfied if $m(w) = -e^{-rw}$, $r > 0$, as in the CARA example given earlier. Note that if $m(w(x_1))$ is increasing and concave in x_1, then so is $v(w(x_1), x_2)$. It is for this reason that Assumption A3' will prove to be sufficient to justify the FOA. However, to use Jewitt's argument, it is evidently necessary to strengthen Assumptions A2 and A4.

Assumption A2' (Concave likelihood-ratio): The marginal distributions satisfy Assumption A2 (MLRP). Moreover, $l^1_{a_1}(x_1|a_1)$ is weakly concave in x_1 for all $x_1 \in [\underline{x}_1, \overline{x}_1]$.

Assumption A4' (Multiplicative rewards): The agent's Bernoulli utility is $-m(w)n(x_2) - c(a_1, a_2)$. Costs, $c(a_1, a_2)$, are strictly increasing and weakly convex in (a_1, a_2). The tasks are weak substitutes; $c_{12}(a_1, a_2) \geq 0$. The rewards functions m and n are strictly negative functions that are strictly increasing and strictly concave on their domain. Finally, m satisfies (17).

The FOA can now be justified in the multiplicative model.

Theorem 3 *Assume the second-best action (a_1, a_2) is interior. Then, given Assumptions A1,A2',A3',A4', A6', and A7, the FOA is valid.*

Proof. Starting from the expression of $EU(a_1, a_2)$ derived in Theorem 2, another round of integration by parts leads to a new expression that depends only on the terms in Assumption A3' (rather than A3" as in Theorem 2). Concavity then obtains if the derivatives of $u(x_1, x_2) = v(w(x_1), x_2)$ have the correct sign. It is required that $u_1, u_2 \geq 0 \geq u_{11}, u_{12}, u_{22}$ and $u_{112}, u_{122} \geq 0 \geq u_{1122}$. Assumptions A2 and A4 together implies that $u_{11} \leq 0$ (as $m(w(x_1))$ is concave in x_1). Given the multiplicative nature of u, it then follows that the second set of inequalities is also satisfied. Thus, the agent's utility is concave. ∎

5.2 Characterizing the optimal action

Assume that $v(w, x_2)$ takes the form in (13). To begin, fix some interior a_1 that the principal might like to induce. To begin, focus on cases where a_1 is accompanied by an interior a_2. In this case, $M(a_1)$ and $M'(a_1)$ are characterized completely by the local incentive compatibility constraints that $EU_1 = EU_2 = 0$, with

$$M'(a_1) = \frac{c_1(a_1, a_2)}{N(a_2)},\; M(a_1) = \frac{c_2(a_1, a_2)}{N'(a_2)}.$$

Next, note that

$$\begin{aligned}
\frac{\partial}{\partial a_2}\left(\frac{c_1(a_1, a_2)}{N(a_2)}\right) &= \frac{c_{12}(a_1, a_2)N(a_2) - c_1(a_1, a_2)N'(a_2)}{N(a_2)^2} > 0 \\
\frac{\partial}{\partial a_2}\left(\frac{c_2(a_1, a_2)}{N'(a_2)}\right) &= \frac{c_{22}(a_1, a_2)N'(a_2) - c_2(a_1, a_2)N''(a_2)}{N'(a_2)^2} < 0
\end{aligned}$$

since $N(a_2), N''(a_2) > 0 > N'(a_2)$. It follows that $M(a_1)$ and $M'(a_1)$ are inversely related, as described in the introduction.

The above discussion makes use of the incentive compatibility constraints only. Once the participation constraint is taken into account as well, it turns out that not all a_2 can be implemented, given a_1. Thus, the first problem is to characterize the "feasible set".

Note that $EU_2 \geq 0$ is equivalent to

$$M(a_1) \leq \frac{c_2(a_1, a_2)}{N'(a_2)}, \tag{18}$$

where $M(a_1)$ and $N(a_2)$ are defined in (14) and (15), respectively, and where $N'(a_2) < 0$. The left hand side depends on the contract, $w(x_1)$, but that dependence is suppressed for notational convenience. The important observation is that if a_2 is interior, then (18) must hold with equality, or the agent would find it optimal to deviate. Thus, M is entirely determined by the a_2 the principal wishes to implement, when a_2 is interior. Note, however, that M is also the only instrument through which the principal can extract rent from the agent, given the action. These observations taken together then suggests that it may not always be possible to make the participation constraint binding. In fact, the participation constraint can be written as

$$M(a_1) \geq \frac{\overline{u} + c(a_1, a_2)}{N(a_2)}. \tag{19}$$

Hence, when a_2 is interior, it is evident that the participation constraint is satisfied if and only if

$$\frac{c_2(a_1, a_2)}{N'(a_2)} N(a_2) - c(a_1, a_2) \geq \overline{u}. \tag{20}$$

Assumption A3, A3', and A3" all imply that $N(a_2)$ is strictly convex in a_2. Simple differentiation then shows that the left hand side is strictly decreasing in a_2. Still holding a_1 fixed, define

$$t(a_1) = \max \left\{ a_2 \in [\underline{a}_2, \overline{a}_2] \,|\, \frac{c_2(a_1, a_2)}{N'(a_2)} N(a_2) - c(a_1, a_2) \geq \overline{u} \right\}$$

as the threshold value of a_2 such that (20) holds for all a_2 below that value. If $t(a_1)$ is interior, (20) is satisfied if and only if $a_2 \in [\underline{a}_2, t(a_1)]$. In other words, only a_2 levels below the threshold $t(a_1)$ can be implemented.[19] If $t(a_1) = \overline{a}_2$, no interior a_2 can be implemented at all. Finally, if $t(a_1) = \underline{a}_2$, then no $a_2 > \underline{a}_2$ can be implemented.[20] As this case is less interesting, it will be ignored in the remainder. That is, assume $t(a_1) > \underline{a}_2$. Note that for any $a_2 \in (\underline{a}_2, t(a_1))$, (20) must be slack, implying that the agent earns more than reservation utility. Now, given that wage costs are continuous on $(\underline{a}_2, t(a_1)]$, it follows that the principal would prefer inducing some $a_2 < t(a_1)$ if $B_2(a_1, t(a_1))$ is sufficiently small (i.e. negative). If this property holds at the second best a_1 (and as long as $a_2 = \underline{a}_2$ is not optimal), the agent will thus earn economic rents. Corollary 1 summarizes.

Corollary 1 *The agent may earn economic rents if $B_2 < 0$.*

Next, I examine how wage costs depend on $a_2 \in [\underline{a}_2, t(a_1)]$. Fix a feasible (a_1, a_2) pair, and let $C(a_1, a_2)$ denote the implementation costs (expected wage costs). I will show that for any interior a_1, $C(a_1, a_2)$ is decreasing in a_2 on $(\underline{a}_2, t(a_1)]$.

To outline the proof, fix an interior (a_1, a_2) pair where the participation constraint is slack, i.e. where $a_2 \in (\underline{a}_2, t(a_1))$. Formulate the cost-minimization problem which derives the cheapest contract that induces (a_1, a_2) subject to the participation constraint and the incentive compatibility constraints that $EU_1 = EU_2 = 0$. Since the participation constraint is slack, the Envelope Theorem implies that a marginal change in a_2 causes $C(a_1, a_2)$ to change by $-(\mu_1 EU_{12} + \mu_2 EU_{22})$, where $EU_{12}, EU_{22} < 0$ and $\mu_1 > 0 > \mu_2$.[21] Thus, costs decrease if μ_1 is small compared to μ_2. To this end, the proof of the result establishes that

$$\mu_1 c_1(a_1, a_2) + \mu_2 c_2(a_1, a_2) < 0$$

in the multiplicative model, whenever the participation constraint is slack.[22] This inequality bounds μ_1 relative to μ_2. Of course, the relative magnitude of EU_{12} and EU_{22} are also relevant for signing $\mu_1 EU_{12} + \mu_2 EU_{22}$. To close the proof, I thus assume

[19] To implement $\overline{a}_2$, it must hold that $EU_2 \geq 0$, which requires (18) to hold. However, the weak inequality does not alter the conclusion that (20) is also necessary to induce $\overline{a}_2$.

[20] It may still be possible to implement $\underline{a}_2$ since (18) need not hold in this case.

[21] Recall from Section 5.1 that $\mu_1 > 0$ holds even when (a_1, a_2) does not coincide with the optimal pair. Since the participation constraint is slack, $\lambda = 0$. Then, (16) would be violated for some x_1 unless $\mu_2 < 0$.

[22] This conclusion is reached by modifying Jewitt's (1988) famously elegant proof that $\mu_1 > 0$ in the single-task setting.

that

$$\frac{\partial}{\partial a_2}\left(\frac{c_2(a_1,a_2)}{c_1(a_1,a_2)}\frac{N(a_2)}{N'(a_2)}\right) \leq 0 \tag{21}$$

for all (a_1, a_2). This is a mild assumption. It is satisfied if e.g. $c_{22} > 0$ or $c_{12} \leq 0$ is numerically sufficiently large. More interestingly, it is also satisfied if $N(a_2)$ is log-convex, which in turn is ensured if $G^2(x_2|a_2)$ or its antiderivative (the counterpart to (10)) is log-convex in a_2. Recall the fact that Assumptions A3, A3', and A3" are satisfied if both G^1 and G^2 have that property.

Proposition 1 *Assume utility from rewards are multiplicative, (21) holds, and that the assumptions in one of Theorems 1, 2, or 3 hold. Assume wages are interior regardless of which feasible (a_1, a_2) pair the principal seeks to implement.*[23] *Then, for any $a_1 \in (\underline{a}_1, \overline{a}_1)$ for which $t(a_1) > \underline{a}_2$, $C(a_1, a_2)$ is strictly decreasing in a_2 on $(\underline{a}_2, t(a_1)]$.*

Proof. In the Appendix. ■

To make the next point, assume that $B_2 = 0$, i.e. the principal does not directly care about a_2. Proposition 1 then signifies that for any (interior) a_1, the optimal a_2 to induce is $a_2 = t(a_1)$, which I will assume to be interior. The same conclusion of course holds if $B_2 \geq 0$. Consequently, the participation constraint is binding regardless of which a_1 the principal seeks to implement. Now, (20) is strictly decreasing in both a_1 and a_2, from which it follows that $t'(a_1) < 0$. In words, if the principal desires the agent to work harder on the job, then it is optimal to at the same time induce the agent to work less hard in the pursuit of private rewards.

Given a_1 is to be implemented, the agent's total costs is thus $c(a_1, t(a_1))$. Although the a_1 and $t(a_1)$ move in opposite directions, it can nevertheless be shown that

$$\frac{\partial c(a_1, t(a_1))}{\partial a_1} \geq 0 \tag{22}$$

given the assumption that (21) holds.[24] That is, the agent's total cost of effort increases when he is induced to work harder, even though he works less intensively on the other task.

[23] Then, the FOA can be used to determine the cost-minimizing contract for any interior (a_1, a_2). The contract takes the form in (8) or (16).

[24] The assumption can be used to bound $t'(a_1)$. Specifically, $t'(a_1) \geq -\frac{c_1(a_1,t(a_1))}{c_2(a_1,t(a_1))}$, which in turn yields (22).

Compare two a_1 levels, a_1'' and a_1', with $a_1'' > a_1'$. Let w'' and w' denote the contract that optimally implements a_1'' and a_1', respectively. Since the participation constraint binds in either case, but costs are higher when a_1'' is implemented, it follows that utility from both types of rewards must be greater under a_1'' than under a_1'. That is,

$$M(a_1''|w'')N(t(a_1'')) \geq M(a_1'|w')N(t(a_1')).$$

However, since $t(a_1'') < t(a_1')$, the agent's private returns are worth less when $a_1 = a_1''$. Hence, he must be compensated with higher rewards at work. Thus, $M(a_1''|w'') > M(a_1'|w')$. Most likely, such a conclusion would appear obvious to any outside observation of the principal-agent relationship, even if he was not cognizant of the endogenous nature of a_2.

However, the next observation may at first appear more surprising. To induce an interior a_1, it is necessary that $EU_1 = 0$, or

$$M'(a_1|w) = \frac{c_1(a_1, a_2)}{N(a_2)},$$

where the right hand side should be evaluated at $a_2 = t(a_1)$. To illustrate the main point in the simplest possible way, assume that marginal costs, $c_1(a_1, a_2)$, are constant. Then,

$$M'(a_1''|w'') = \frac{c_1}{N(t(a_1''))} < \frac{c_1}{N(t(a_1'))} = M'(a_1'|w').$$

Hence, an outsider who fails to take into account that a_2 is endogenous would conclude that the marginal return to extra effort is *lower* the harder the agent is induced to work.[25] In other words, it looks as if the agent works harder when given weaker explicit incentives.

6 Implications and applications

This section is devoted to the economic significance of Theorems 1 and 2. Section 6.1 starts with a quick discussion of how the agent would react to a contract that ignores his ability to manipulate a_2. Section 6.2 examines the robustness of a key predictions of the standard moral hazard model to the inclusion of private rewards. Section 6.3

[25] However, if c_1 is sufficiently steep, the more intuitive conclusion that $M'(a_1''|w'') > M'(a_1'|w')$ is obtained.

discusses the relevance of the current results for common agency problems, i.e. moral hazard problems where competing principals contract with the same agent.

6.1 Incentives with private rewards

The standard moral hazard model ignores private rewards. As a thought experiment, imagine instead that the principal acknowledges that the agent receives private rewards, but does not take into account the fact that the contract will affect not only the agent's choice of a_1, but also his choice of a_2. In other words, the principal believes that a_2 is fixed at some specific value, denoted $\widehat{a}_2$. Equivalently, the principal simply thinks of the distribution of private rewards as being fixed and outside the agent's control. Let a_1^* denote the optimal action that the principal would then seek to implement, and assume $a_1^* \in (\underline{a}_1, \overline{a}_1)$. The contract offered by the principal would then take the form in (8), but without the last term. The contract is monotonic. Since the principal believes the agent has no incentive to deviate, the contract must imply that $EU_1(a_1^*, \widehat{a}_2) = 0$.

The question is now how the agent will react to the above contract. The answer lies in recalling that a_1 and a_2 are substitutes whenever the agent faces a monotonic contract. That is, $EU_{12} < 0$. Since it is also the case that $EU_{11} < 0$, the curve along which $EU_1(a_1, a_2) = 0$ is decreasing in (a_1, a_2) space and by design passes through the point $(a_1^*, \widehat{a}_2)$. The agent's best response, assuming it is interior, must lie somewhere along this curve. Only in knife-edge cases will it be true that $EU_2(a_1^*, \widehat{a}_2) = 0$ as well, since the principal in the current thought experiment ignores the agent's incentives to pursue private rewards. Thus, the agent generally moves to another point on the aforementioned curve. It now follows that the agent's best response generically entails either (i) $a_1 > a_1^*$ and $a_2 < \widehat{a}_2$ or (ii) $a_1 < a_1^*$ and $a_2 > \widehat{a}_2$. In other words, the agent will work harder than the principal expects on one task, and less hard on the other task. The principal's assessment of $\widehat{a}_2$ determines which case applies. For instance, if $\widehat{a}_2 = \underline{a}_2$ and $a_1^* \in (\underline{a}_1, \overline{a}_1)$, then the agent will work less hard for the principal than expected. The opposite holds if $\widehat{a}_2 = \overline{a}_2$.

6.2 Interpreting the multipliers

It is well-known that in the standard single-task model, $\mu_1 > 0$ implies that the principal would be better off if the agent (by mistake) works marginally higher than

a_1^* on the task. The same is true in the current model, though it takes slightly more effort to establish it.

Proposition 2 *Assume the second-best action is in the interior. Then, given Assumptions A1–A8, the principal is (weakly) better off if the agent works marginally harder than a_1^*, or*

$$B_1'(a_1^*, a_2^*) - \int w(x_1) g_{a_1}^1(x_1|a_1^*) dx_1 \geq 0. \tag{23}$$

Proof. Given $EU_1(a_1^*, a_2^*) = 0$, the adjoint equation for a_1 reduces to

$$B_1'(a_1^*, a_2^*) - \int w(x_1) g_{a_1}^1(x_1|a_1^*) dx_1 + \mu_1 EU_{11}(a_1^*, a_2^*) + \mu_2 EU_{12}(a_1^*, a_2^*) = 0. \tag{24}$$

Now, substitute (9) into (24), to get

$$\left(B_1'(a_1^*, a_2^*) - \int w(x_1) g_{a_1}^1(x_1|a_1^*) dx_1 \right) + \mu_1 \frac{EU_{11}(a_1^*, a_2^*) EU_{22}(a_1^*, a_2^*) - EU_{12}(a_1^*, a_2^*)^2}{EU_{22}(a_1^*, a_2^*)} = 0.$$

Recall the denominator in the last term is strictly negative. Since the solution is incentive compatible, $EU(a_1, a_2)$ must be concave at (a_1^*, a_2^*). Hence, the numerator must be non-negative. Then, (23) follows from $\mu_1 > 0$. ■

Given the properties of the optimal contract identified in Lemma 2, the last term on the left hand side in (23) is strictly increasing; wage costs are higher when the agent works harder on task a_1. Hence, (23) necessitates that $B_1'(a_1^*, a_2^*) > 0$. In other words, (a_1^*, a_2^*) can be optimal only if the principal's benefit function is increasing in a_1 at that point.

The fact that $\mu_2 < 0$ implies that the principal is better off if the agent's marginal cost on task 2, $c_2(a_1, a_2)$, were to increase. This is not surprising given that tasks are strict substitutes. Other things being equal, a higher marginal cost on task 2 lowers the agent's desire to work on the private task. This is turn increases his return to task 1 (as $EU_{12} < 0$). Thus, it becomes easier, or cheaper, to persuade the agent to work on the task that is directly relevant to the principal.

In hindsight, then, the model's results agree with common intuition. This is a comforting outcome. After all, the standard model – limited as it is – has now been shown to possess some robustness in light of outside random disturbances. The FOA can still be justified, the optimal contract changes in a natural way (see (7)), and the intuitive result that $\mu_1 > 0 > \mu_2$ obtains.

While it may appear unsurprising that $\mu_2 < 0$, it is worthwhile to compare the present setup with private rewards to a situation in which the principal can observe and contract on x_2. For simplicity, think of x_2 as outside income. Technically, the agent is the recipient of x_2, but the principal can make his transfer, $w(x_1, x_2)$ contingent upon it, such that the agent's total income is $w(x_1, x_2) + x_2$. Conceptually, it is useful to "change variables" and instead think of the principal as being the recipient of x_2, in which case the transfer to the agent from the principal is $W(x_1, x_2) = w(x_1, x_2) + x_2$. This is now a more or less standard moral hazard problem, albeit the agent is still multi-tasking. Kirkegaard (2014) shows that the FOA is valid given weaker assumptions than those in Theorem 1. However, it also follows from his analysis that both multipliers are strictly positive. Hence, with private rewards, the negative multiplier is not due to the outside rewards, but rather due entirely to the assumption that the reward is secret.

6.3 Common agency

Common agency refers to a situation in which the agent works on behalf of two (or more) competing principals. Bernheim and Whinston (1986) were first to consider such situations. They assume that every principal observes the same information. Thus, any principal can observe and verify how well the agent performed for other principals. The principals play a non-cooperative game in which they simultaneously offer contracts. Bernheim and Whinston (1986) establish that the equilibrium action is implemented at a total cost that coincides with the total cost that would have obtained if the principals could collude (or merge). As Bernheim and Whinston (1986) explain: "We can always view a principal as constructing his incentive scheme in two steps: he first undoes what all the other principals have offered and then makes an 'aggregate' offer [...]. Clearly, if we are at an equilibrium, each principal must, in this second step, select an aggregate offer that implements the equilibrium action at minimum cost." On the other hand, competition between principals typically distorts the equilibrium action away from the second-best. Note that Bernheim and Whinston (1986) do not prove that an equilibrium of the game exists in general, although existence is proven in two special cases.

Evidently, the assumption that principals share the same information is often unrealistic. The preceding analysis of principal-agent problems with private rewards may now be of some use. Think of a_i as the effort devoted to principal i's job, $i = 1, 2$. Thus, assume principal i cares directly only about a_i. In contrast to Bernheim and Whinston (1986), assume that principal i cannot observe x_j, nor how much the agent is paid by principal j or for that matter the contract offered by principal j, $j \neq i$. In this case, principal i's contract can be written $w^i(x_i)$, since the wage schedule he offers can depend only on x_i. The wage the agent receives from principal j, $w^j(x_j)$, is from principal i's point of view akin to a private reward.

Consider principal 1's problem. Imagine first that $w^2(x_2)$ is continuous and strictly increasing, and let $x^2(w^2)$ denote its inverse. Since the function w^2 transforms x_2, the relevant joint distribution changes from (1) to

$$H(x_1, w^2|a_1, a_2) = G^1(x_1|a_1)G^2(x^2(w^2)|a_2).$$

As long as the original distribution function, $F(x_1, x_2|a_1, a_2) = G^1(x_1|a_1)G^2(x_2|a_2)$, satisfies Assumption A3, $H(x_1, w^2|a_1, a_2)$ must also satisfy Assumption A3. A similar property holds for Assumption A2. Assumption A7 must be modified slightly; it is necessary that any solution to the relaxed problem involves only interior wages for all continuous and strictly increasing $w^2(x_2)$ functions. Given this adaptation, however, Theorem 1 applies. In other words, if $w^2(x_2)$ is continuous and strictly increasing, then *it is a best response for principal 1 to also offer a continuous and strictly increasing wage schedule*, $w^1(x_1)$, at least as long as an interior action is to be implemented.[26] Given the difficulties faced by Bernheim and Whinston (1986) in an informationally simpler set-up, equilibrium existence is not pursued here and is in any event tangential to the main point of the paper. In the remaining discussion it is simply assumed that an equilibrium exists.

Although the above observation may appear intuitive, it is in fact a new result. The closest existing result that I am aware of is due to Holmström and Milgrom (1988). Their model is built on their earlier seminal contribution, Holmström and Milgrom (1987), in which they consider a dynamic moral hazard problem where the agent can adjust his effort over time. Assuming effort costs are monetary and that the agent has constant absolute risk aversion (i.e. exponential utility), their problem can

[26] Mirroring the argument following Theorem 1, the action must be componentwise strictly greater than $(\underline{a}_1, \underline{a}_2)$ whenever $c_1(\underline{a}_1, a_2) = c_2(a_1, \underline{a}_2) = 0$ for all a_1, a_2.

be reduced to a static moral hazard problem in which the agent controls the means of jointly normally distributed variables and where the principal is restricted to using linear contracts. Holmström and Milgrom (1988) show that this logic extends to a situation with competing principals. That is, if one principal uses a linear contract, then it is a best response for the other to do so as well.

Holmström and Milgrom's (1988) model – sometimes referred to as the Linear-Exponential-Normal (LEN) model – has the advantage that is it quite tractable. Thus, Holmström and Milgrom (1988) are able to compare equilibrium under different informational assumptions. They use the term "disjoint observations" to refer to the case where principal i observes only x_i (private rewards) and the term "joint observations" to the case where the pair (x_1, x_2) is observed by both principals. A main result is that principals may prefer disjoint observations to joint observations.[27]

Kirkegaard (2014) provides several examples that the equilibrium predictions of the LEN model with a single principal are not robust, thus demonstrating that the LEN model's tractability comes at a price. In the common agency setting, the LEN model's predictions differ in an additional way from the model of the current paper, as established next.

When signals are independent in the LEN model, it is easy to see that the equilibrium action is implemented in a cost-minimizing manner. That is, given independence, Bernheim and Whinston's (1986) result on joint observations extends to disjoint observations in the LEN model. The underlying reason is that independence together with linear contracts and exponential utility imply so much "separability" that nothing is gained from collusion.

However, it generally remains the case that disjoint observations distorts the equilibrium action. An exception is when $c_{12}(a_1, a_2) = 0$ for all (a_1, a_2), in which case the "second best" is implemented under disjoint observations.[28] Here, second best refers to the action that the principals would implement if they collude and share information (or together offer a single (linear) contract based on (x_1, x_2)). This is in part due to the assumption that effort costs are monetary in the LEN model. Given independence, CARA utility, and $c(a_1, a_2) = c^1(a_1) + c^2(a_2)$ (or $c_{12}(a_1, a_2) = 0$), the

[27] Maier and Ottaviani (2009) reconsiders this issue in a LEN model where the agent's action is one-dimensional, meaning that the single choice variable affects the distribution of both x_1 and x_2.

[28] In the LEN model, a_i is the mean of x_i, $i = 1, 2$.

agent then seeks to maximize

$$-\int\left(-e^{-r\left(w^1(x_1)-c^1(a_1)\right)}\right)g^1(x_1|a_1)dx_1\times\int\left(-e^{-r\left(w^2(x_2)-c^2(a_2)\right)}\right)g^2(x_2|a_2)dx_2, \quad (25)$$

under disjoint observations, where $r>0$ is the coefficient of absolute risk aversion. Here, the agent maximizes his utility by maximizing each of the two terms separately. Assuming the agent participates, principal i's contract can thus in no way influence how hard the agent works for principal j. Consequently, there is no "externality" between the principals, and no distortion away from the second best. Under joint observations, however, principal i can influence a_j. As Holmström and Milgrom (1988) note, "principal 1 can [...] offer to insure the agent partly against the excess risk that principal 2 is trying to impose on the agent to provide better incentives. Such side-contracting destroys the socially optimal contract that principal 2 would otherwise have designed." Thus, when $c_{12}=0$, *(i)* the equilibrium action under disjoint observations coincides with the second best action, and *(ii)* it is implemented at second best costs. In contrast, the equilibrium action under joint observations is distorted. Hence, although the latter is also implemented at minimum cost, as established by Bernheim and Whinston's (1986), the principals prefer the former.[29]

In the model in the current paper, however, it is generally the case that the equilibrium action under disjoint observations is both *(i)* different from the second best and *(ii)* implemented at higher than minimum costs, even when $c_{12}=0$. The reason is that the current model lacks some of the separability that is so plentiful in the LEN model. For concreteness, and to make the point the most forcefully, assume the agent has constant absolute risk aversion, or $v(w)=-e^{-rw}$, $r>0$, but that costs are not monetary, contrary to the LEN model. In particular, when $c(a_1,a_2)=c^1(a_1)+c^2(a_2)$, the agent maximizes

$$-\int\left(-e^{-rw^1(x_1)}\right)g^1(x_1|a_1)dx_1\times\int\left(-e^{-rw^2(x_2)}\right)g^2(x_2|a_2)dx_2-c^1(a_1)-c^2(a_2). \quad (26)$$

Hence, a change in $w^1(x_1)$ may lead the agent to change a_2 as well. In other words,

[29] There is an indeterminancy in how the surplus is shared among the principals, however. Given contracts of the form $w^i(x_i)=\beta^i+\alpha^i x^i$, what matters for the agent's participation and incentive constraints are α^1,α^2, and $\beta^1+\beta^2$. Thus, β^1 and β^2 are indeterminate. However, there must be an equilibrium where both principals are better off under disjoint observations than under joint observations.

principal 1's contract impacts how hard the agent works for principal 2. Consequently, under disjoint observation, the equilibrium action is distorted away from the second best. Moreover, the equilibrium action is implemented at higher than minimum costs. To see this, note first that Kirkegaard (2014) shows that the FOA is valid with a single principal who observes (x_1, x_2). Thus, the FOA is also valid with two colluding and information sharing principals. The FOA contract is then described by

$$\frac{1}{v'(w(x_1,x_2))} = \lambda + \mu_1 \frac{g^1_{a_1}(x_1|a_1)}{g^1(x_1|a_1)} + \mu_2 \frac{g^2_{a_2}(x_2|a_2)}{g^2(x_2|a_2)},$$

where $\mu_1, \mu_2 > 0$. Since $v'(w) = re^{-rw}$, it follows that

$$\frac{e^{rw(x_1,x_2)}}{r} = e^{\ln\left(\lambda + \mu_1 \frac{g^1_{a_1}(x_1|a_1)}{g^1(x_1|a_1)} + \mu_2 \frac{g^2_{a_2}(x_2|a_2)}{g^2(x_2|a_2)}\right)}. \tag{27}$$

In comparison, with disjoint observations one would obtain

$$\frac{1}{v'(w^1(x_1) + w^2(x_2))} = \frac{e^{r\left(w^1(x_1)+w^2(x_2)\right)}}{r}.$$

However, (27) can never take this form when $\mu_1, \mu_2 > 0$. Thus, it is not possible for the two contracts under disjoint observations to mimic the cost minimizing contract obtained when the two principals collude and share information.

Table 1 contrasts the main conclusions from the LEN model with the main conclusions from the current model, under the assumption that signals are independent and $c_{12} = 0$. In summary, in the current model the equilibrium action is distorted away from the second best under both joint and disjoint observations. Under joint observation, however, the equilibrium action is implemented at minimum costs. This is not the case under disjoint observations. Thus, it is not obvious that Holmström and Milgrom's (1988) LEN result that disjoint observations are preferable when e.g. $c_{12} = 0$ hold in the current model. Maier and Ottaviani (2009) explain Holmström and Milgrom's (1988) result by noting that: "Given that under private contracting [i.e. disjoint observations] each principal has access to a performance measure about the task about which she cares, the common agency distortion is avoided altogether and the second-best outcome results." However, this argument is invalid in the current model, for two reasons. First, effort costs are modeled differently; compare (25) and (26). Second, contracts are not restricted to be linear in any information regime; disjoint contracts typically do not aggregate in a way that mimics the cost minimizing contract.

An example suffices to demonstrate that Holmström and Milgrom's (1988) result can be overturned. Thus, consider a situation where the agent's marginal costs are very small compared to the principals' marginal benefits. If the difference is sufficiently large, it seems natural to conjecture that the equilibrium action, under either information regime, will be $(a_1, a_2) = (\overline{a}_1, \overline{a}_2)$. In this case, the cost of implementation is higher under disjoint observation than under joint observation, but the equilibrium action and thus the direct benefit to the principals are the same in either case. Hence, they prefer joint observation in this example.

I next outline a proof that $(\overline{a}_1, \overline{a}_2)$ is indeed an equilibrium action under disjoint observation. Evidently, for $(\overline{a}_1, \overline{a}_2)$ to be an equilibrium action, it must hold that $EU_i(\overline{a}_1, \overline{a}_2) \geq 0$, $i = 1, 2$. In the relaxed cost-minimization problem, it thus holds that $\mu_1, \mu_2 \geq 0$. Assume that the agent exhibits constant absolute risk aversion, such that the multiplicative model applies. Under disjoint observation, it is then easy to see that $EU_i(\overline{a}_1, \overline{a}_2) > 0$ cannot solve the relaxed problem. The reason is that principal i's contract would be flat (as the multiplier on the $EU_i(\overline{a}_1, \overline{a}_2) \geq 0$ constraint is zero), which would in turn violate $EU_i(\overline{a}_1, \overline{a}_2) \geq 0$. Thus, $EU_i(\overline{a}_1, \overline{a}_2) = 0$, precisely as when an interior action is implemented.[30] Thus, using the argument in e.g. Lemma 1, the agent has no incentive to deviate from $(\overline{a}_1, \overline{a}_2)$. This leaves only the possibility that the principals will deviate. It is clear that principal 1 does not have an incentive to deviate and induce some $a_2 < \overline{a}_2$. This follows from the fact that $t(a_1)$ is decreasing whenever it is defined, meaning that since $\overline{a}_2$is feasible when $\overline{a}_1$ is induced it is feasible regardless of which a_1 is induced. The results in Section 5.2 then proves that $\overline{a}_2$ is the optimal companion to any a_1. The fact that principal 1 does not wish to deviate and induce $a_1 < \overline{a}_1$ follows from the assumption that his benefit function is very steep in a_1. After all, wage costs are continuous in a_1, since the constraints $EU_1 = EU_2 = 0$ are the same regardless of which $a_1 > \underline{a}_1$ is implemented.[31] Thus, when principal 1's benefit function is sufficiently steep in a_1, there is no incentive to sacrifice benefits for a marginal decrease in wage costs. By a symmetric argument, principal 2 has no

[30] Recall that the last term in (8) is constant in the multiplicative model. Outside the multiplicative model positive multipliers might produce more than one solution to (8), thereby complicating the argument.

[31] If $c_1(\underline{a}_1, a_2) = 0$, then it is impossible to implement $a_1 = \underline{a}_1$.

incentive to deviate either.

	Joint observations		Disjoint observations	
	Eq. action	Cost of eq. action	Eq. action	Cost of eq. action
LEN model	$\neq$ second-best	cost-minimizing	second-best	cost-minimizing
Current model	$\neq$ second-best	cost-minimizing	$\neq$ second-best	>cost-minimizing

Table 1: Comparing models with independence and $c_{12} = 0$.

7 Conclusion

The purpose of the current paper is to extend the canonical principal-agent model to allow the agent the possibility to pursue private rewards. These private rewards may be monetary or non-monetary; as described in the introduction, it is easy to think of examples of both.

Since the agent is multi-tasking in this setting, a justification of the FOA necessitates an understanding of the basic moral hazard problem with multi-tasking. However, multi-tasking has been largely ignored in the literature (the LEN model being an exception) until very recently. The justification of the FOA presented here thus builds upon Kirkegaard's (2014) analysis. As explained there, the main technical cost of allowing multi-tasking is that tasks must be assumed to be stochastically independent. With this restriction in place, however, the current paper establishes additional conditions under which Kirkegaard's (2014) justifications extends to private rewards. Once the required assumptions on the technology have been made, as identified in Kirkegaard (2014), the economically significant assumptions are that the agent perceives outcomes and tasks to be substitutes and that his absolute risk aversion over labor income is decreasing in the private reward. It should be stressed that these assumptions appear to be rather mild.

The model of private rewards presented here is fairly simple, and thus abstracts away from a few potentially important complications. As just mentioned, a key assumption is that rewards are independent. However, it is not inconceivable that for example the gifts parents bestow on their children depend on the job held by the latter. Strictly speaking, the model does not allow the distribution of private rewards to be a direct function of the contract. However, the distribution could depend on the type of profession the agent is employed in, much in the same way that the outside option is likely to be a function of the agent's profession or level of education.[32]

[32] The decision of which profession to enter or how much education to acquire is not modelled.

References

Ábrahám, Á., Koehne, S. and N. Pavoni (2011): "On the first-order approach in principal-agent models with hidden borrowing and lending," *Journal of Economic Theory*, 146: 1331-1361.

Athey, S. (2002): "Monotone Comparative Statics under Uncertainty," *Quarterly Journal of Economics*, 117: 187-223.

Bénabou, R. and J. Tirole (2003): "Intrinsic and Extrinsic Motivation," *Review of Economic Studies*, 70: 289-520.

Bernheim, B.D. and M.D. Whinston (1986): "Common Agency," *Econometrica*, 54 (4): 923-942.

Conlon, J.R. (2009): "Two new Conditions Supporting the First-Order Approach to Multisignal Principal-Agent Problems," *Econometrica*, 77 (1): 249-278.

Englmaier, F. and S. Leider (2012): "Contractual and Organizational Structure with Reciprocal Agents," *American Economic Journal: Microeconomics*, 4 (2): 146-183.

Holmström, B. and P. Milgrom (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55 (2): 303-328.

Holmström, B. and P. Milgrom (1988): "Common Agency and Exclusive Dealing," mimeo.

Holmström, B. and P. Milgrom (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, 24-52.

Jewitt, I. (1988): "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica*, 56 (5): 1177-1190.

Laffont, J-J. and D. Martimort (2002): The Theory of Incentives: The Principal-Agent Model, Princeton University Press.

Kirkegaard, R. (2014): "A Unifying Approach to Incentive Compatibility in Moral Hazard Problems", mimeo (October 2014).

Kőszegi, B. (2014): "Behavioral Contract Theory," *Journal of Economic Literature*, 52 (4): 1075-1118.

Maier, N. and M. Ottaviani (2009): "Information Sharing in Common Agency: When is Transparency Good?" *Journal of the European Economic Association*, 7 (1): 162-187.

Rogerson, W.P. (1985): "The First-Order Approach to Principal-Agent Problems," *Econometrica*, 53 (6): 1357-1367.

Appendix

Proof of Proposition 1. Since a_1 is interior, the condition that $EU_1 = 0$ must hold. In other words, $M'(a_1)N(a_2) - c_1(a_1, a_2) = 0$, or

$$\int m(w(x_1))\mu_1 l^1_{a_1}(x_1|a_1)g^1(x_1|a_1)dx_1 N(a_2) = \mu_1 c_1(a_1, a_2).$$

Borrowing a trick from Jewitt (1988), solve (16) for $\mu_1 l^1_{a_1}$, and rewrite the above to yield

$$\int m(w(x_1))\left(\frac{1}{m'(w(x_1))} - \mu_2 N'(a_2)\right) g^1(x_1|a_1)dx_1 = \mu_1 c_1(a_1, a_2)$$

whenever the participation constraint is slack ($\lambda = 0$), as it must be when $a_2 \in (\underline{a}_2, t(a_1))$. Whenever a_2 is also interior, $EU_2 = 0$, or $N'(a_2) = \frac{c_2(a_1,a_2)}{M(a_1)}$, thus producing

$$\begin{aligned} \int \frac{m(w(x_1))}{m'(w(x_1))} g^1(x_1|a_1)dx_1 &= \mu_1 c_1(a_1, a_2) + \mu_2 \int m(w(x_1))\frac{c_2(a_1, a_2)}{M(a_1)} g^1(x_1|a_1)dx_1 \\ &= \mu_1 c_1(a_1, a_2) + \mu_2 c_2(a_1, a_2). \end{aligned}$$

Since m is negative and increasing, the left hand side is negative. In summary,

$$\mu_2 c_2(a_1, a_2) < -\mu_1 c_1(a_1, a_2) < 0.$$

Assume now that $a_2 \in (\underline{a}_2, t(a_1))$. As explained in the text, a small increase in a_2 then reduces implementation costs by

$$\begin{aligned} \mu_1 EU_{12} + \mu_2 EU_{22} &= \mu_1 \left[M'(a_1)N'(a_2) - c_{12}\right] + \mu_2 \left[M(a_1)N''(a_2) - c_{22}\right] \\ &= \mu_1 \left[\frac{c_1}{N(a_2)}N'(a_2) - c_{12}\right] + \mu_2 \left[\frac{c_2}{N'(a_2)}N''(a_2) - c_{22}\right] \\ &= \mu_1 c_1 \left[\frac{N'(a_2)}{N(a_2)} - \frac{c_{12}}{c_1}\right] + \mu_2 c_2 \left[\frac{N''(a_2)}{N'(a_2)} - \frac{c_{22}}{c_2}\right] \\ &> \mu_1 c_1 \left[\frac{N'(a_2)}{N(a_2)} - \frac{c_{12}}{c_1}\right] - \mu_1 c_1 \left[\frac{N''(a_2)}{N'(a_2)} - \frac{c_{22}}{c_2}\right] \\ &= \mu_1 c_1(a_1, a_2)\frac{N'(a_2)}{N(a_2)}\frac{c_2(a_1, a_2)}{c_1(a_1, a_2)}\frac{\partial}{\partial a_2}\left(\frac{c_2(a_1, a_2)}{c_1(a_1, a_2)}\frac{N(a_2)}{N'(a_2)}\right) \\ &> 0, \end{aligned}$$

where the second equality uses $EU_1 = EU_2 = 0$. The first inequality uses the bound on $\mu_2 c_2$ derived earlier, combined with the fact that the term in the last bracket is negative. The second inequality invokes (21). Thus, costs are strictly decreasing on $a_2 \in (\underline{a}_2, t(a_1))$. The solution to the stated cost-minimization problem is continuous in a_2, and hence $a_2 = t(a_1)$ is the cheapest way of inducing a_1 on $(\underline{a}_2, t(a_1)]$.[33] ■

[33] The solution to the stated cost-minimization problem may even over-estimate the cost at $a_2 = t(a_1)$. The reason is that if $t(a_1) = \overline{a}_2$, the constraint that $EU_2 = 0$ can be replaced by the weaker $EU_2 \geq 0$. A similar observation applies at $a_2 = \underline{a}_2$, which is why it is not claimed that $C(a_1, a_2)$ is maximized at $a_2 = \underline{a}_2$.