Moral Hazard and the Spanning Condition without the First-Order Approach*

René Kirkegaard†

June 2016

Abstract

The “spanning condition” describes a situation where the agent’s effort determines the weights placed on two distinct technologies. Sufficient conditions are known under which the first-order approach (FOA) is valid when the spanning condition holds. In this paper, a complete solution to the problem is provided. Thus, the problem is solved even in cases where the FOA is not valid. The solution has two main steps. The first step fully characterizes the set of actions that can be implemented. The second step establishes that the optimal contract can be found by applying the FOA to the newly identified set of feasible actions. In short, a modified, or two-step, FOA solves the problem. This procedure correctly solves a famous counterexample due to Mirrlees in which the standard FOA is not valid. A much simpler and economically more compelling counterexample is provided. Moreover, the model provides a tractable environment in which it is possible to study comparative statics even when the FOA is not justified. An example involving threshold monitoring illustrates. Finally, the solution is generalized to the case where the agent determines the weights placed on an arbitrary number of technologies. The agent’s action is multi-dimensional in this case.

JEL Classification Numbers: D82, D86

Keywords: First-Order Approach, Moral Hazard, Principal-Agent Models, Spanning Condition

*I would like to thank the Canada Research Chairs programme and the Social Sciences and Humanities Research Council of Canada for funding this research. I am grateful to four anonymous referees for constructive comments.

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

A general solution to the moral hazard problem remains elusive. Assuming the agent’s action is continuous, it is common in applications to utilize the first-order approach (FOA). The FOA relies on the assumption that the agent will have no incentive to deviate from the action targeted by the principal as long as the agent’s utility achieves a stationary point at said action. Stated differently, only “local” incentive compatibility is deemed relevant. A still-active literature has developed several sets of conditions under which the FOA is valid.\(^1\) Nevertheless, it is well-understood that the FOA is not always valid. That is, sometimes the agent must be prevented from deviating to more remote actions; “non-local” incentive compatibility constraints may be binding. Another common approach simply assumes that the agent’s effort is either high or low, in which case the incentive compatibility problem is easily dealt with. Both approaches are more or less explicitly designed to minimize any complications that can arise from the general incentive compatibility problem. Certainly, both models owe their popularity to this very fact.

In contrast, the express purpose of the current paper is to tackle the non-local incentive compatibility constraints head on. Thus, the paper is related to a smaller literature, from Grossman and Hart (1983) to Kadan and Swinkels (2013), which attempts to obtain economic insights without relying on the FOA. I deliberately focus on a particularly tractable moral-hazard environment. The model is just rich enough that the FOA is not always valid, yet simple enough that a complete solution is still possible. The idea is that scrutinizing a simple model may yield significant economic insights into the role of the non-local constraints.\(^2\)

Consider an agent whose action, \(a\), is continuous but unverifiable. The contract between the principal and agent is instead based on the verifiable vector of signals, \(x\). Given action \(a\), the distribution of \(x\) is denoted \(F(x|a)\). The simplifying assumption imposed in this paper is that the distribution satisfies the spanning condition, i.e. that \(F(x|a)\) can be written as

\[
F(x|a) = p(a)G(x) + (1 - p(a))H(x),
\]


\(^2\)Araujo and Moreira (2001) take a different approach, as they attempt to solve general moral hazard problems. While technically very sophisticated, the level of abstraction is such that economic insights are harder to come by.
where \( p(a) \in [0, 1] \) for all \( a \in [a, \bar{a}] \) and \( G \) and \( H \) are non-identical distribution functions. While this model is certainly too specialized to capture all principal-agent relationships, it should be stressed that it does have a compelling interpretation. For instance, \( p(a) \) could be the proportion of time the parts-supplier (the agent) spends using the new and advanced technology \( G \) rather than the less reliable but more user-friendly old technology, \( H \).

Distributions of this form have been studied extensively. Grossman and Hart (1983) first introduced the spanning condition. Without fail, additional assumptions are imposed on either the curvature of \( p(a) \) or on the relationship between \( G \) and \( H \), or both. These extra assumptions are then typically used to justify the FOA.\(^3\) No complete analysis has been offered, or attempted, until now.

It can perhaps be argued that the focus on the FOA is counterproductive. The fact of the matter is that the spanning condition embodies a lot of structure on its own. Thus, it is tempting to believe that a full solution to the problem is possible, without further regularity assumptions. In this endeavour, the FOA is a distraction.

Hence, no regularity conditions are imposed in this paper beyond differentiability. A complete solution to the problem is nevertheless obtained. To achieve the solution, a simple insight from a companion paper is exploited. Specifically, Kirkegaard (2015) observes that the local incentive compatibility constraint by itself imposes some structure on the contract. He then uses this observation to pursue new justifications of the FOA in general moral hazard models. However, he does not attempt to move beyond the FOA. In this paper, in contrast, the structure in the local incentive compatibility constraint is combined with the specific structure in (1) to obtain a full solution to the moral hazard problem under the spanning condition.

It is useful to think of the solution as consisting of two easy steps. In the first step, the set of implementable actions is characterized in a simple and intuitive manner. For instance, if \( p(a) \) is strictly increasing and costs are normalized to be linear in \( a \), then \( a' \) is implementable if and only if \( a' \) is on the concave hull of \( p(a) \).\(^4\) The set of implementable actions is independent of \( G \) and \( H \).

---

\(^3\)For instance, Sinclair-Desgagné (1994, 2009) points out that the FOA is valid if \( p(a) \) is concave and \( \frac{g(x)}{h(x)} \) is nondecreasing, where \( g \) and \( h \) are the densities of \( G \) and \( H \), respectively. The latter assumption implies the monotone likelihood-ratio property (MLRP). Ke (2013a) and Jung and Kim (2015) show that MLRP is not necessary.

\(^4\)The concave hull is the lowest concave function that is always above \( p(a) \). For future reference, the convex hull is the highest convex function that is always below \( p(a) \). See Rockafellar (1970).
The second step describes how to find a solution among the actions in the feasible set. For actions that are interior and implementable, the local incentive compatibility constraint turns out to be sufficient. Thus, the FOA can be applied to this set of actions in order to identify the most profitable interior action. An equally simple procedure determines whether it would be even better to induce a corner action. In summary, a solution to the problem is obtained through what is essentially a modified FOA. As explained in the following, the modified FOA is useful for several reasons.

The simplicity of the modified FOA means that the model is highly tractable even when the FOA is not valid. Thus, the spanning condition is suited to explore optimal contracting and comparative statics outside the confines of the FOA. To illustrate the value of the model in this regard, an example involving what I term threshold monitoring is examined. Threshold monitoring refers to an extreme monitoring technology that enables the principal to observe and verify actions below a threshold. The higher the threshold, the higher the monitoring costs to the principal. The value of monitoring is in part that it effectively eliminates a set of non-local incentive compatibility constraints, thus making it possible to induce actions that are not otherwise implementable. This effect is absent when the FOA is valid, since non-local constraints are redundant in that case. Hence, the comparative statics are different depending on whether the FOA is valid or not. When the FOA is not valid, a reduction in the cost of the technology makes the principal invest more heavily in monitoring but he may at the same time induce a lower, hitherto unimplementable, action. Thus, the level of monitoring and the level of effort may move in opposite directions.

It is well known that the FOA is not always valid. Mirrlees (1999) offers a famous example to illustrate how the FOA may fail by identifying the wrong action as being optimal. In their textbook, Bolton and Dewatripont (2005, p. 148) remark that: “This example is admittedly abstract, but this is the only one to our knowledge that addresses the technical issue.” I will establish that the modified FOA correctly solves Mirrlees’ (1999) example. A more straightforward example of how the FOA may fail is provided. The example is conceptually simpler and unlike Mirrlees’ example it takes the form of a textbook principal-agent problem. A complementary example shows that the FOA may also fail by identifying the correct optimal action but an incorrect contract. In the current model, this can happen only if a corner action is optimal. Finally, the modified FOA makes short shrift of the leading example in Araujo and Moreira’s (2001) rather technical paper.
Ke (2013a) and Jung and Kim (2015) prove that the FOA is valid if $p(a)$ is increasing and concave. I allow $p(a)$ to be non-concave. Nevertheless, concavity of $p(a)$ may seem natural. One interpretation is that the agent’s production exhibits decreasing returns to scale, in the sense that increasing effort increases the probability of $x$ being drawn from $G$ instead of $H$ at a decreasing rate. Thus, it is valid to ask why non-concave $p(a)$ are interesting. There are at least three reasons.

The first and perhaps most fundamental reason is that decreasing returns to scale rules out some economic phenomena. For instance, there may be an element of learning-by-doing as the agent first starts working. In this case, $p(a)$ is first convex and it only later becomes concave as decreasing returns set in. See Section 4 for more examples in this vein. Second, and echoing a previous point, the simplicity and tractability of the model means that it can be used as a laboratory of sorts to study optimal contracts and comparative statics when the FOA is not valid. In order to exploit the model for this purpose, however, $p(a)$ must necessarily be assumed to be non-concave. Third, the examples alluded to above demonstrate that by explicitly tackling non-concave $p(a)$, the approach in the current paper unifies examples that are important in the literature and makes it possible to develop simpler examples.

The paper concludes by examining a generalization of (1) in which the agent determines the weights placed on an arbitrary number, $m + 1$, of technologies. In this setting, then, the agent’s action is $m$-dimensional. Note that a “full dimensionality” assumption is imposed here; regardless of the first $i$ ($i \leq m - 1$) weights, the remaining $m - i$ weights are never predetermined. Associated with each dimension is a local incentive compatibility constraint. As before, the structure implicit in the contract is then sufficient to solve the problem. The modified FOA remains valid. An implication is that the principal-agent problem is easily solved if the number of outcomes is finite and the agent determines the probability with which each occurs.

2 Model

The agent takes a costly and unverifiable action, $a$, which belongs to some compact interval, $[a, \bar{a}]$. Think of the agent’s action as being the cost of effort he is willing to incur. This normalization implies that the cost function is $c(a) = a$. The contract is

---

5This particular normalization is chosen for ease of comparison with the existing literature, most of which assume that $p(a)$ has curvature. In Section 5, I consider another normalization, used by
based on an $n$-dimensional random vector of signals, where $\mathbf{x} = (x_1, ..., x_n)$ denotes a realization. Let $w(\mathbf{x})$ denote the wage stipulated by the contract if the realization is $\mathbf{x}$. It is for notational simplicity only that the wage is assumed to be deterministic; stochastic wages can be incorporated, at the cost of more notation.\footnote{Hart and Holmström (1987), in which the probability weights are linear, but where the cost function has curvature. The latter normalization is better suited to multi-dimensional extensions.} Critically, utility is additively separable, meaning that the agent’s utility is $v(w(\mathbf{x})) - c$ if the signal realization is $\mathbf{x}$, where $v$ describes the utility from income. Assume that $v$ is continuous and strictly increasing on a convex domain. Finally, $v$ is unbounded above and/or below. These assumptions on $v$ are imposed merely to ensure the existence of a contract which satisfies local incentive compatibility (see Section 3). The agent is assumed to be an expected utility maximizer, but at no point are any assumptions imposed on the agent’s or principal’s risk preferences. Thus, the agent is allowed to be e.g. loss averse (with an S-shaped $v$ function).

Given the action $a$, the distribution function takes the form in (1). A generalization is in Section 5. Recall the assumption that $G$ and $H$ are distinct. Assume that $p(a)$ is continuously differentiable. Given the interpretation of the model in the introduction, the most meaningful economic assumption is that $p(a)$ is monotonic. Thus, as is standard in the literature, assume that $p'(a) > 0$ for all $a \in (a, \bar{a}]$. This assumption is for expositional convenience; the case where $p(a)$ is non-monotonic is not that much more difficult and is treated in brief below.

No other assumptions on the primitives are imposed. For instance, $p(a)$ may be concave only locally, or not at all, and $G$ and $H$ may cross, as would be the case if $H$ is a mean-preserving spread over $G$. Indeed, $G$ and $H$ need not be continuous nor have the same support. However, note that $F(\mathbf{x}|a)$ has the same support for all $a$ for which $p(a) \in (0, 1)$. The reason is that there is a non-zero probability that $\mathbf{x}$ is drawn from $G$ or $H$ whenever $p(a) \in (0, 1)$. Hence the support of $F(\mathbf{x}|a)$ is simply the union of the supports of $G$ and $H$. For future reference, let $F_a(\mathbf{x}|a) = p'(a)(G(\mathbf{x}) - H(\mathbf{x}))$ denote the derivative of $F(\mathbf{x}|a)$ with respect to $a$.

Throughout, only contracts for which the agent’s expected utility exists and is bounded will be considered. Technically, it may be possible to implement any action by providing the agent with unbounded utility, but presumably such contracts would
be prohibitively costly for the principal.\footnote{Kadan and Swinkels (2013) examine what they term “boundedly implementable” actions, i.e. actions that can be implemented with a bounded contract.}

Given the contract, the agent’s expected utility is

\[
EU(a) = \int v(w(x))dF(x|a) - a
\]  

(2)

when he takes action \(a\). In order for some action \(a^*\) to be implemented, it must be in the agent’s self-interest to pick that particular action, or

\[
EU(a^*) \geq EU(a) \text{ for all } a \in [a, \bar{a}],
\]

(G-IC\(_{a^*}\))

in which case, following Kirkegaard (2015), the contract \(w(x)\) is said to be globally incentive compatible. If \(a^* \in (a, \bar{a})\), it is necessary that \(EU(a)\) attains a stationary point at \(a^*\), or

\[
\int v(w(x))dF_a(x|a) - 1 = 0.
\]

(L-IC\(_{a^*}\))

The above condition will be referred to as the local incentive compatibility constraint. Hence, any contract for which \(EU'(a^*) = 0\) is termed L-IC\(_{a^*}\) and any contract that satisfies \(EU(a^*) \geq EU(a)\) for all \(a \in [a, \bar{a}]\) is G-IC\(_{a^*}\).

Exploiting the structure inherent in the spanning condition, expected utility can be written as

\[
EU(a) = p(a)\beta + \gamma - a
\]  

(3)

where

\[
\beta = \int v(w(x))d(G(x) - H(x)) , \quad \gamma = \int v(w(x))dH(x).
\]

(4)

Recall that although \(\beta\) and \(\gamma\) are fixed from the agent’s point of view, they are in fact determined by the principal. Hence, the principal adjusts \(\beta\) and \(\gamma\) depending on which action he wants to induce. From (3), \(\beta\) can be used to manipulate the agent’s incentives, while \(\gamma\) can be adjusted to ensure participation. From (4), \(\beta\) is the expected bonus (measured in utils) from drawing \(x\) from \(G\) rather than \(H\). Since \(a\) determines the probability of these two complementary events, the agent thus responds to changes in \(\beta\).
3  A modified first-order approach

The analysis is in two main steps. The first is a full characterization of the set of actions for which a G-IC contract exists. This question takes as given that the agent has signed the contract in the first place. In other words, the participation constraint is initially ignored. Likewise, the principal’s objective function is irrelevant to the question of whether he can manipulate the agent to take a certain action. The second main step consists of identifying a procedure that correctly derives the optimal contract on the feasible set of actions. The implementation of this step requires adding the principal’s objective function, the participation constraint, and whatever other constraints may be present in that specific contracting environment.

3.1  The feasible set of actions

As in Kirkegaard (2015), a key step is to explore the link between L-IC and G-IC. Using the notation in (3), L-IC is

\[ p'(a^*) \beta - 1 = 0 \]  

whenever \( a^* \in (\underline{a}, \bar{a}) \). Thus, if the principal intends to induce \( a^* \) he must at a minimum adjust \( \beta \) to ensure that (5) is satisfied. Since \( p'(a^*) > 0 \), \( \beta \) must therefore be set equal to the strictly positive value \( \frac{1}{p(a^*)} \). Given this value,

\[ EU(a^*) - EU(a) = \frac{[p(a^*) + (a - a^*)p'(a^*)] - p(a)}{p'(a^*)}. \]  

(6)

Note that the bracketed term is the tangent line to \( p(a) \) through \( a^* \). Let \( A^C \) denote the set of actions in \((\underline{a}, \bar{a})\) for which \( p(a) \) coincides with its concave hull. By definition of the concave hull, \( a^* \in A^C \) if and only if the tangent line at \( a^* \) lies everywhere above \( p(a) \), so (6) is non-negative for any \( a \).

**Proposition 1** Assume that \( p'(a) > 0 \) for all \( a \in (\underline{a}, \bar{a}) \). Then, there exists a G-IC\( a^* \) contract (that yields bounded utility) if and only if \( a^* \in A^C \cup \{\underline{a}, \bar{a}\} \).

**Proof.** Assume \( a^* \in (\underline{a}, \bar{a}) \) and \( a^* \notin A^C \). If there is a G-IC\( a^* \) contract, then that contract must necessarily be L-IC\( a^* \), and so (6) applies. However, since \( a^* \notin A^C \), there is some \( a \in (\underline{a}, \bar{a}) \) for which (6) is strictly negative, which contradicts G-IC\( a^* \).
For the other direction, assume \( a^* \in A^C \). Since \( G \) and \( H \) are distinct, there is some \( x' \) for which \( G(x') \neq H(x') \). Now, construct a step contract that yields utility \( v_0 \) if \( x \leq x' \), and utility \( v_1 \) otherwise. Then, \( \beta = (v_0 - v_1) (G(x') - H(x')) \). Recall that L-IC\(_{a^*}\) requires \( \beta = \frac{1}{p'(a^*)} \). Thus, since utility is assumed to be continuous and unbounded above and/or below, there exists a \((v_0, v_1)\) pair that satisfies L-IC\(_{a^*}\). Since \( a^* \in A^C \), (6) is everywhere non-negative. Hence, the contract is G-IC\(_{a^*}\).

Consider next actions at the corners, starting with \( a^* = \bar{a} \). Here, \( EU(\bar{a}) - EU(a) = (p(\bar{a}) - p(a)) \beta - (\bar{a} - a) \) is positive for all \( a < \bar{a} \) if \( \beta \) is sufficiently large. Stated differently, a step contract with a large \( \beta \) is G-IC\(_{\bar{a}}\). Again, \( \beta \) can be made arbitrarily large by manipulating \((v_0, v_1)\). As an aside, note that \( EU'(\bar{a}) \) may be strictly positive. At the other corner, \( \underline{a} \) can be implemented with a flat wage even if \( p'(\underline{a}) = 0 \).

Figure 1 provides two examples. In panel (i), \( p(a) = \frac{1}{20} + 2(a^2 - a^3) \), \( a \in [0, 0.65] \). This function is convex on \([0, \frac{1}{3}]\) and concave on \([\frac{1}{3}, 0.65] \). However, \( A^C = [\frac{1}{2}, 0.65] \). That is, the concave hull of \( p(a) \) coincides with the dashed line on \([0, \frac{1}{2}]\) and with the function \( p(a) \) thereafter. Note that for \( a \) to be on the concave hull of \( p \) it is necessary but not sufficient that \( p \) is concave at that point. The function in panel (ii) is \( p(a) = \frac{2}{3} \sqrt{a} - (a - \frac{2}{3})^3 - \frac{1}{4} \), \( a \in [0, 1] \). This function is concave-convex-concave. The convex part in the middle cannot be on the concave hull. Indeed, in this case \( A^C = (0, 0.057] \cup [0.834, 1) \). Thus, \( A^C \) need not be convex.

![Figure 1: The concave hull of two non-concave functions.](image-url)
3.2 The optimal contract and the modified FOA

Proposition 1 shows that the spanning condition allows a succinct formulation of the “feasible set” of implementable actions, \( A^C \cup \{a, \bar{a}\} \). The feasible set is closed but not necessarily convex. Transitioning to the question of how to optimally implement any given feasible action, it should be clear from the proof of Proposition 1 that L-IC\(_a^*\) is not only necessary but also sufficient for G-IC\(_{a^*}\), for any \( a^* \in A^C \).

**Proposition 2** Assume that \( p'(a) > 0 \) for all \( a \in (a, \bar{a}] \). If \( a^* \in A^C \) then any L-IC\(_a^*\) contract is G-IC\(_{a^*}\).

**Proof.** Given \( p'(a) > 0 \), (6) is everywhere positive if \( a^* \in A^C \). ■

In their discrete-action model, Grossman and Hart (1983) allow more than one incentive compatibility constraints to bind even when the spanning conditions is imposed. However, Proposition 2 implies that for any implementable action all but the local incentive compatibility constraint are redundant in the continuous model. Thus, the only role of non-local constraints is that they are responsible for limiting the set of implementable actions in the first place. Stated differently, it is enough to worry only about L-IC\(_a^*\) when thinking about how to optimally induce some \( a^* \in A^C \). However, things are more complicated when \( a \) or \( \bar{a} \) are to be induced. For this reason, it is useful to treat the two parts of the feasible set, \( A^C \) and \( \{a, \bar{a}\} \), separately.

Propositions 1 and 2 suggest a modified, or two-step, FOA for dealing with interior actions. In the first step, \( A^C \subseteq (a, \bar{a}) \) is identified. In the second step, the FOA is applied to this set, i.e. with the constraint that \( a \in A^C \). This step requires adding the participation constraint (if applicable) to the problem. Any number of exogenous constraints on the contracting environment could also be permitted at this stage. For instance, the contract may need to be monotonic or to satisfy a limited liability constraint. Having thus obtained the constraint set, it is only at this point that it becomes necessary to specify the principal’s objective function, including his risk preferences and how his payoff depends on wages, the outcome.

---

\(^8\)\(A^C\) is a half-open interval or the union thereof in both panels of Figure 1. If \( p(a) \) is concave, then \( A^C = (a, \bar{a}) \) is open. However, if \( p(a) \) is convex near \( a \) and \( \bar{a} \), but concave in the middle, then \( A^C \) will consist of a closed interval of \( a \)’s on the concave part. In short, \( A^C \) can take many forms, yet \( A^C \cup \{a, \bar{a}\} \) is always closed.

\(^9\)Hermalin and Katz (1991) use tools from convex analysis to characterize the set of implementable actions in a model with a finite set of actions and a finite set of outcomes. However, their analysis does not reveal when L-IC is sufficient for G-IC.
For instance, if the principal is an expected utility maximizer with a general Bernoulli utility function of the form $B(w, x, a)$, his maximization problem can be written as

$$\max_{w(x), a \in A_C} \int B(w(x), x, a)dF(x|a),$$

subject to L-IC$_a$, the participation constraint, and possibly a set of additional constraints. As in Grossman and Hart (1983), this problem can be solved by first finding the optimal contract for any $a \in A_C$, and then next selecting the $a$ in $A_C$ that yields the highest payoff to the principal. Note that a solution to the second part is not guaranteed to exist since $A_C$ might be an open set. This is, however, just an artifact of having split the closed feasible set into two parts, one of which might be open.

The solution to the above problem evidently depends on the constraint set and the principal’s objective function. However, even without fully specifying these elements, it is possible to meaningfully compare the optimal contract with the optimal contract that would be obtained from uncritically applying the FOA. Still ignoring the issue of implementing actions at the corner, the only difference between the standard FOA and the modified FOA is that the former takes the feasible set to be $(a, \bar{a})$ rather than $A_C \subseteq (a, \bar{a})$ in (7). Thus, if the standard FOA suggests an optimal action in $A_C$, then the solution to the modified FOA would be exactly the same. Conversely, if the standard FOA suggests an action not in $A_C \cup \{a, \bar{a}\}$ then it has failed to correctly solve the problem. In this case, however, the modified FOA produces a contract that has the same general structure as if the FOA was valid, in the sense that L-IC is the only relevant incentive compatibility constraint; see Examples 2 and 3 in Section 4.

Consider next the implementation of $a$ and $\bar{a}$. Specifying the local incentive compatibility constraint as $EU'(a) = 0$ is potentially too restrictive when $a \in \{a, \bar{a}\}$. For instance, the local condition for implementing $a$ is the weaker condition that $EU'(a) \leq 0$. Moreover, non-local constraints may also matter for the implementation of corner actions. However, it turns out that the continuum of incentive compatibility constraints can again be summarized by one lone condition. Thus, as with actions in $A_C$, the optimal contract that implements $a$ or $\bar{a}$ can be obtained by maximizing the principal’s objective subject to a single incentive compatibility constraint.

---

10The action $a$ is indirectly relevant since it determines the distribution over outcomes and wages. However, it may also enter the principal’s utility function directly. This is perhaps particularly realistic in cases where $a$ is observable by the principal but not verifiable, i.e. not contractible.
I begin with a closer look at implementing $a$. Let $a^c = \inf A^C$ if $A^C$ is non-empty and let $a^c = \pi$ otherwise. First, $EU'(a) \leq 0$ is necessary for G-IC$_a$. However, if $a = a^c$, then $EU'(a) \leq 0$ is also sufficient for G-IC$_a$, as proven below. Consider next the possibility that $a < a^c$. Then, G-IC$_a$ obviously necessitates that $EU(a) \geq EU(a^c)$, such that there is no incentive to pick $a^c$ over $a$. However, it turns out that $EU(a) \geq EU(a^c)$ is in fact sufficient for G-IC$_a$. In particular, $EU(a) \geq EU(a^c)$ implies $EU'(a) \leq 0$ when $a < a^c$.

To implement $\pi$, the relevant counterpart to $a^c$ is $\bar{a}^c = \sup A^C$ when $A^C$ is non-empty and $\bar{a}^c = a$ otherwise. In panel (i) of Figure 1, $a^c = \frac{1}{2}$ and $\bar{a}^c = \bar{a}$. Note that contrary to actions in $A^C$, non-local incentive compatibility constraints may have an impact on the optimal implementation of actions at the corners.

**Proposition 3** Assume that $p'(a) > 0$ for all $a \in (a, \pi]$. Then, it is possible to implement the boundary actions, as follows:

1. If $a^c = a$ then $EU'(a) \leq 0$ is necessary and sufficient for G-IC$_a$. If $a^c > a$ then $EU(a) \geq EU(a^c)$ is necessary and sufficient for G-IC$_a$.

2. If $\bar{a}^c = \bar{a}$ then $EU'(\bar{a}) \geq 0$ is necessary and sufficient for G-IC$_\pi$. If $\bar{a}^c < \bar{a}$ then $EU(\bar{a}) \geq EU(\bar{a}^c)$ is necessary and sufficient for G-IC$_\pi$.

**Proof.** Necessity is obvious. For sufficiency in the first part of the proposition, consider first the “no-gap” case, $a^c = a$. Here, the slope of $p(a)$ coincides with the slope of its concave hull at $a$. As in the proof of Proposition 1, a modification of (6) then establishes that $EU'(a) \leq 0$ is sufficient for G-IC$_a$. However, this is not necessarily true in the “gap” case, where $a^c > a$. Note that

$$EU(a) - EU(a) = (a - a) \left[ 1 - \frac{p(a) - p(a)}{a - a} \beta \right].$$

Hence, $EU(a) \geq EU(a^c)$ implies that the term in brackets must be non-negative when $a = a^c$. If $\beta$ is negative, then the term in brackets is positive for all $a$, or $EU(a) \geq EU(a)$ for all $a$. That is, the contract is G-IC$_a$. If $\beta$ is positive, then the term in brackets is minimized at $a = a^c$. This follows by definition of the concave hull, since the line from $(a, p(a))$ to $(a^c, p(a^c))$ is steeper than the line from $(a, p(a))$ to any other point on $p(\cdot)$. Hence, $EU(a) \geq EU(a^c)$ implies $EU(a) \geq EU(a)$ for all $a \in [a, \pi]$, i.e. G-IC$_a$. The proof of the second part of the proposition is analogous. ■
A short summary of the analysis follows. Proposition 1 fully characterizes the feasible set of actions. Propositions 2 and 3 together identify the single incentive compatibility constraint that matters for implementing any given feasible action. Thus, it is conceptually straightforward to derive the optimal contract that implements any specific feasible action and then to subsequently select the best action to implement.

Ke (2013a) and Jung and Kim (2015) prove that the FOA is valid if $p(a)$ is increasing and concave, even without the monotone likelihood ratio property. More formally, Ke (2013a) defines the FOA to be valid if $G$-IC is replaced by the three-part constraint that (i) $EU'(a^*) = 0$ if $a^* \in (a, \bar{a})$, (ii) $EU'(a^*) \leq 0$ if $a^* = a$, and (iii) $EU'(a^*) \geq 0$ if $a^* = \bar{a}$, and the solution to this relaxed problem also solves the unrelaxed problem. The result in Ke (2013a) and Jung and Kim (2015) is thus a corollary of Propositions 1–3, since $A^C = (a, \bar{a})$ and $a^c = a$ and $\bar{a}^c = \bar{a}$ in this case.

In fact, Propositions 1 and 2 prove that the first part of this three-part constraint generalizes under the modification that the domain must be changed to $A^C$. If it also the case that $a^c = a$ and $\bar{a}^c = \bar{a}$, then Proposition 3 proves that the other two parts of the constraint are sufficient for the actions at the corner. Thus, the FOA is valid when the domain of actions is taken to be $A^C \cup \{a, \bar{a}\}$. The example in panel (ii) of Figure 1 can thus be solved in this manner.

Proposition 3 characterizes when the last two parts of the previous three-part constraint are in fact sufficient for G-IC. In particular, this requires $a^c = a$ and $\bar{a}^c = \bar{a}$, respectively. Assume instead that $\bar{a}^c < \bar{a}$, say. Then, the optimal contract that induces $\bar{a}$ is the same as the optimal contract that induces $\bar{a}$ in a two-action model where the agent only has two actions at his disposal, $a^c$ and $\bar{a}$. In this case, then, the FOA may identify, correctly or incorrectly, the optimal action to be $\bar{a}$, but misspecify the optimal contract that induces $a$; see Example 5 in Section 4.

### 3.3 Discussion and extensions

From (6) it immediately follows that if $a^* \in A^C$, then the agent is indifferent between $a^*$ and any $a$ for which

$$[p(a^*) + (a - a^*) p'(a^*)] - p(a) = 0.$$

Assume $a^* = \frac{1}{2}$ in the example in panel (i) in Figure 1. Then, the agent would be indifferent between $a^* = \frac{1}{2}$ and $a = 0$. In other words, it is possible to identify exactly
which non-local incentive compatibility constraints are binding. Thus, the model leads to very specific insights into the nature of the global incentive compatibility problem. Consequently, it is feasible to examine comparative statics even when the FOA is not valid; see Example 4 in Section 4. Note that the set of binding non-local incentive compatibility constraints for action $a^*$ are completely independent of the contract offered to implement $a^*$ (as long as it satisfies L-IC$_{a^*}$). In a more general model, it is not unlikely that the binding constraint will depend on the contract.

The characterization of the set of implementable actions depends only on $p(a)$. Apart from the minimal assumption that $G$ and $H$ are distinct, the set of implementable actions is therefore independent of $G$ and $H$. It is also independent of the agent’s risk preferences and the principal’s utility function.

Finally, the assumption that $p(a)$ is monotonic seems justified on economic grounds. However, it is possible to allow $p(a)$ to be non-monotonic. First, note that the argument following (6) remains valid if $p'(a^*) > 0$ even if $p'(a) < 0$ for some $a \neq a^*$. That is, $a^*$ can be implemented, and L-IC$_{a^*}$ is sufficient, if and only if $a^*$ is on the concave hull of $p(a^*)$, given $p'(a^*) > 0$. Secondly, consider some $a^{**} \in (\underline{a}, \overline{a})$ for which $p'(a^{**}) < 0$. Following the same logic as in (6), $a^{**}$ can be implemented, and L-IC$_{a^{**}}$ is sufficient, if and only if $a^{**}$ is on the convex hull of $p(a^{**})$.\footnote{This is easily seen by multiplying both numerator and denominator in (6) by $-1$.} Thus, the set of implementable interior actions can be obtained by piecing together the sets of implementable actions with $p'(\cdot) > 0$ and $p'(\cdot) < 0$, respectively.\footnote{Note that if $a^* \in (\underline{a}, \overline{a})$ and $p'(a^*) = 0$ then no L-IC$_{a^*}$ contract exist, as can be seen from (5). Similarly, if $p(a^*) = p(a')$, then $a^*$ cannot be implemented if $a' < a^*$ because it would be cheaper for the agent to pick $a'$ rather than $a^*$. Consequently, once $p(a)$ is allowed to be non-monotonic, it is no longer necessarily the case that $\overline{a}$ can be implemented.}

4 Examples

The examples in this section fall into two groups. The first group demonstrate the usefulness of the characterization in Section 3. Example 1 considers a problem in Araujo and Moreira (2001). Propositions 1 and 3 greatly simplify that particular problem. Example 2 revisits Mirrlees’ (1999) example in which the FOA is not valid. I carefully describe how the environment in Mirrlees’ example can be reformulated to fit an environment in which the spanning condition holds. Thus, it is possible to use the previous analysis to explain why the FOA fails in his example. Likewise, the
modified FOA correctly solves the problem. This example is not entirely trivial in its original form, as alluded to in the quote by Bolton and Dewatripont (2005) in the introduction. Example 3 therefore provides a much simpler example that makes the exact same point that Mirrlees sought to demonstrate. In these examples the FOA fails because it identifies a non-implementable action as being optimal. Example 4 is a continuation of Example 3, meant to illustrate that comparative statics are possible under the spanning condition even when the FOA is not valid. Example 5 provides another example in which the FOA fails. Here, the FOA identifies the correct optimal action but not the optimal contract that induces that action.

The last two examples provide more perspective on the model. Example 6 illustrates why \( p(a) \) might realistically not always be concave. Example 7 observes that several distributions that are used in the recent literature satisfy the spanning condition. It is explained how those examples are related to the previous analysis.

Before moving on to the examples, note that any setting in which there are two outcomes but a continuum of actions is a special case of (1). Specifically, a two-outcome model is obtained by assuming that \( G \) and \( H \) are degenerate distributions, with all mass concentrated at different points. In fact, Examples 1-5, including Mirrlees’ example, can all be thought of as two-outcome models.

**Example 1 (Araujo and Moreira (2001)):** Araujo and Moreira (2001) propose a general Lagrangian approach to the moral hazard problem that applies when the FOA is not valid. Their leading example is the following. There are two states, where state 1 is the bad state and state 2 is the good state. The agent picks an effort level, \( e \), with \( e \in [0,0.9] \). With effort \( e \), the probability of the good state is \( q(e) = e^3 \). The cost of effort is \( c(e) = e^2 \). To reparameterize the model, let \( a \equiv c(e) = e^2 \) and \( p(a) = q(c^{-1}(a)) = a^{3/2}, a \in [0,0.81] \). Note that \( p(a) \) is increasing and convex. Thus, \( A^C \) is empty. In other words, no interior action can be implemented (Proposition 1). However, the boundary actions can be implemented, and the only relevant incentive compatibility constraint is that the desired action be preferable to the action on the opposite end of the support (Proposition 3). Consequently, this example essentially reduces to the textbook model with two outcomes and two actions and is therefore trivial to solve once a participation constraint is added; see Example 5. In contrast, to use their general approach to solve the example, Araujo and Moreira (2001) (having added assumptions on \( v(w) \) and on the principal’s payoff) construct an algorithm in Mathematica and use this to solve 20 non-linear systems of equations. As expected,
they find the optimal action is at a corner. While their method is obviously powerful, using it on their leading example is overkill and obscures the deeper intuition that interior actions are not even implementable. Ke (2013b) proposes another method to solve this problem. Though his method is simpler than that used by Araujo and Moreira (2001), it remains more complicated than the method suggested above.

**Example 2 (Mirrlees (1999))**: Mirrlees (1999) examines the following example. Consider an agent with payoff function

\[
U(w, z) = we^{-(z+1)^2} + e^{-(z-1)^2}
\]

and a principal with payoff function

\[
V(w, z) = -(z - 1)^2 - (w - 2)^2.
\]

The principal controls the scalar \(w\), whereas \(z\) is the agent’s choice variable. There is no participation constraint in this example; it is as if the agent is forced to work for the principal. Mirrlees’ (1999) approach to the problem is described first. Then, the problem is reformulated in a way that allows the insights developed in Section 3 to be used in order to solve the problem.

![Figure 2: Mirrlees’ example.](image)

The dot in Figure 2 indicates the principal’s bliss point. The dashed curves surrounding it represents contour lines or indifference curves. Next, consider the
agent. Given some $w$ stipulated by the principal, the agent’s first-order condition is satisfied when

$$w = \frac{1 - z}{z + 1}e^{4z}. \quad (10)$$

The curve marked L-IC$_z$ in Figure 2 describes this first-order condition. It is denoted L-IC$_z$ because (10) identifies the (unique) value of $w$ that makes the agent’s first-order condition hold at the target value of $z$. As in Mirrlees (1999), I focus on $w \geq 0$ which in turn limits the solutions to (10) to the range $z \in (-1, 1]$. Note that $w = 0$ implements $z = 1$, and that the ensuing point, (1, 0), is closer to the principal’s bliss point than any $(z, w)$ for which $w < 0$. Hence, there is no value in considering $w < 0$.

Mirrlees observes that for $w$ in an intermediate range, $w \in [0.344, 2.903]$, there are three values of $z$ which solve (10). The critical implication is that the agent’s first-order condition may not identify a global maximum in this case. Of the three candidates, the middle value always identifies a local minimum, while the other two values identify local maxima. For instance, at $w = 1$, any $z \in \{-0.957, 0, 0.957\}$ solve (10). In this particular case, with $w = 1$, (8) is symmetric around zero. Hence, the agent is indifferent between $z = -0.957$ and $z = 0.957$. When $w > 1$, the first term in (8) is weighted more heavily. It becomes more important to increase $- (z + 1)^2$ instead of $-(z - 1)^2$. For this reason, $z < 0$ is uniquely optimal when $w > 1$ and $z > 0$ is uniquely optimal when $w < 1$. The fat part of L-IC$_z$ in Figure 2 identifies the agent’s best response to any $w$, thus describing the true feasible set. The elements of Figure 2 described thus far essentially reproduce Mirrlees’ (1999) Figure 2.

The FOA maximizes (9) subject to (10). The solution is found where the principal’s indifference curve is tangent to L-IC$_z$. Here, $w = 1.99$ and $z = 0.895$. However, the agent’s best response to $w = 1.99$ is in fact not 0.895 but instead close to $-1$. Hence, the FOA fails. By inspecting Figure 2, it is evident that the feasible point that is closest to the bliss point is $(z, w) = (0.957, 1)$. Thus, the correct solution has the principal offering $w = 1$ and recommending to the agent that he selects $z = 0.957$.

Note that the above process used to find the correct solution can be recast as what is essentially a modified FOA. First, the set of implementable $z$ is identified. This rules out that $z$ is in the intermediate range $(-0.957, 0.957)$. Then, (9) is maximized subject only to L-IC$_z$ and the restriction that $z$ belongs to the feasible set. What

---

*I am grateful to an anonymous referee for suggesting this argument.*
is perhaps less obvious is how the identification of the feasible set is related to the approach used in Section 3 to find the feasible set. The first step towards explaining this part consists of reformulating the problem, as follows.

It may be helpful to think of the setup as a special environment with two outcomes, where, for some reason, the wage in one state is exogenously fixed at 0. The principal controls the “bonus” \( w \) if the other state materializes. The agent’s action is \( z \in \mathbb{R} \). Think of \( e^{-(z+1)^2} \) as the probability of the state in which a bonus is paid out, and think of \( -e^{-(z-1)^2} \) as the cost function. As in Example 1, reparameterize the problem by thinking of \( a = -e^{-(z-1)^2} \) as the agent’s choice variable. Thus, costs are normalized to be linear in the agent’s action. Note that \( a \in [-1, 0) \) regardless of \( z \). This function is marked \( a(z) \) in Figure 2. Given \( z \) is no greater than one, the inverse function is \( z(a) = 1 - \sqrt{-\ln(-a)} \). Thus, the probability that the bonus-paying state occurs can now be expressed as a function of \( a \), with

\[
p(a) = e^{-(z(a)+1)^2} = e^{-\left(2-\sqrt{-\ln(-a)}\right)^2}.
\]

Given the reformulation of the problem, the agent’s expected utility now takes the familiar form \( wp(a) - a \). The analysis in Section 3 proves that the properties of \( p(a) \) are important. First, \( p(a) \) is increasing on \([-1, -e^{-4}) \) and decreasing on \((-e^{-4}, 0) \). Thus, actions in the latter interval cannot be implemented with a positive bonus. To see the relationship with Mirrlees’ analysis, note that \( z(-e^{-4}) = -1 \). Second, \( p(a) \) is concave on \([-1, -0.9178) \), convex on \((-0.9178, -0.0542) \), and finally concave again on \((-0.0542, 0) \). In other words, \( p(a) \) has the same qualitative features as the function depicted in panel (\( ii \)) of Figure 1. Thus, as in that figure, \( A_C \) is the union of two disjoint interval, with \( A_C = (-1, -0.9982] \cup [-0.0217, 0) \).\(^{14}\) Given the restriction to positive bonuses, only actions in

\[
(-1, -0.9982] \cup [-0.0217, -e^{-4})
\]

can be implemented. Using the definition of \( a(z) \) above, a link to Mirrlees’ analysis can be established by noting that \( a(0.957) = -0.9982 \) and that \( a(-0.957) = -0.0217 \); see Figure 2. Likewise, it holds that \( p'(0.9982) = p'(-0.0217) = 1 \). This too is unsurprising in hindsight. The reason is that L-IC\(_a\) requires \( w = \frac{1}{p'(a)} \), which means

\(^{14}\)The reason that \( p(a) \) is not plotted and reference is instead made to Figure 1 is that the intervals in \( A_C \) are so small that they are hard to make out in a figure.
that $w = 1$ is required to satisfy L-IC at $a \in \{-0.9982, -0.0217\}$. Thus, as in Mirrlees’ analysis, $w = 1$ is a pivotal value.

In the reformulated problem, the principal’s payoff is given by $\ln (-a) - (w - 2)^2$. One way to conclude the analysis would be to substitute $w = \frac{1}{p'(a)}$ into the objective function and then maximize over $A^C$. This procedure is used in Example 3, below. Here, I opt for a graphical analysis that is closer in spirit to Mirrlees’ approach. Specifically, Figure 3 plots $\text{L-IC}_a$, or $w = \frac{1}{p'(a)}$. The fat parts of the curve identify the set of implementable actions, $A^C$. The principal’s bliss point is at $(-1, 2)$, which is also indicated in the figure along with a few indifference curves. As in Figure 2, the point of tangency identifies an action that is not implementable. Again, the FOA fails. Instead, the optimum is at a corner of the feasible set, where $a = -0.9982$ and $w = 1$. As mentioned already, $a = -0.9982$ translates into $z = 0.957$, which is the solution identified by Mirrlees.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{A reformulation of Mirrlees’ problem.}
\end{figure}

**Example 3 (A simplified counterexample):** There are two outcomes. The probability of the good outcome is $p(a)$, with $p'(a) > 0$. Let $a^*$ denote the action that the principal is seeking to induce. Given a contract that successfully implements $a^*$, let $v_1(a^*)$ be the agent’s utility from wages if the outcome is bad and $v_2(a^*)$ be his utility if the outcome is good. The outcomes are worth $x_1$ and $x_2$ to the risk-neutral principal, respectively. The agent’s reservation utility is $\bar{u}$. The participation
constraint and the L-IC constraint for an interior \( a^* \) require that

\[
v_1(a^*) + p(a)(v_2(a^*) - v_1(a^*)) - a = \overline{u} \\
p'(a)(v_2(a^*) - v_1(a^*)) - 1 = 0
\]

are satisfied at \( a = a^* \). The two constraints thus make it possible to solve directly for the unique candidates for \( v_1(a^*) \) and \( v_2(a^*) \), yielding

\[
v_1(a^*) = \overline{u} + a^* - \frac{p(a^*)}{p'(a^*)}, \quad v_2(a^*) = \overline{u} + a^* + \frac{1 - p(a^*)}{p'(a^*)}.
\] (11)

With these utilities in hand, wages can be obtained by inverting the agent’s utility function. Thus, the candidate contract stipulates wage \( w_1 = v^{-1}(v_1(a^*)) \) if the outcome is bad and wage \( w_2 = v^{-1}(v_2(a^*)) \) if the outcome is good. The caveat is of course that G-IC may be violated, or, stated differently, that \( a^* \) is not implementable at all. However, under the proviso that \( a^* \) is implementable and interior, the principal’s expected payoff is

\[
\pi(a^*) = (1 - p(a^*))(x_1 - v^{-1}(v_1(a^*))) + p(a^*)(x_2 - v^{-1}(v_2(a^*)�)
\] (12)

The standard FOA prescribes maximizing (12). The problem is that this procedure may identify an action that is not implementable.

For concreteness, assume \( p(a) = a + \frac{1}{2}(a^2 - a^3), a \in [0, 2] \). As in the example in panel (i) in Figure 1, \( p(a) \) is convex when \( a < \frac{1}{3} \) and concave when \( a > \frac{1}{3} \). However, the relevant set is \( A^C \), which is \( A^C = \left[ \frac{1}{3}, 2 \right] \). Thus, the set of implementable actions is \( \{0\} \cup \left[ \frac{1}{3}, \frac{2}{3} \right] \). Assume \( \overline{u} = 2 \) and \( v(w) = \sqrt{w} \), in which case (11) yields

\[
v_1(a^*) = \frac{4 + 4a^* - 5(a^*)^2 - 2(a^*)^3}{2 + 2a^* - 3(a^*)^2}, \quad v_2(a^*) = \frac{6 + 4a^* - 5(a^*)^2 - 2(a^*)^3}{2 + 2a^* - 3(a^*)^2}.
\] (13)

Finally, assume that \( x_1 = 5 \) and \( x_2 = 9.4 \). Figure 4 plots \( \pi(a^*) \), again under the pretense that all actions are implementable.\(^{16}\) As in Figures 2 and 3, the fat parts of

\(^{15}\)The participation constraint must bind. Otherwise, the principal could profitably lower both \( v_1(a^*) \) and \( v_2(a^*) \) without affecting the L-IC constraint.

\(^{16}\)In this example, (12) is valid if \( a^* \in \{0, \frac{2}{3} \} \). By Proposition 3, \( a^* = \frac{2}{3} \) can be implemented with any contract for which \( EU'(\frac{2}{3}) \geq 0 \). However, \( EU'(\frac{2}{3}) > 0 \) cannot be optimal because such a contract unnecessarily imposes more risk on the agent (\( v_2 - v_1 \) is larger). Hence, \( EU'(\frac{2}{3}) = 0 \), and so the optimal contract is given by (11), which in turn yields (12). At the other
the curve indicate those actions that are in actual fact implementable.

The function in Figure 4 attains its maximum at $a^{**} = 0.464$. However, $a^{**}$ is not implementable because $a^{**} \notin A^C$. Thus, the FOA fails. Recalling that the feasible set is $\{0\} \cup [\frac{1}{2}, \frac{2}{3}]$, it is clear from Figure 4 that the optimal action to induce is $a^* = \frac{1}{2}$.

The example concludes with two remarks. First, in this example there is a particularly simple way of establishing that the contract implicit in (11) fails to implement the intended action, $a^*$, when $a^* \in (0, \frac{1}{2})$. From (13), $v_1(a^*)$ is inversely U-shaped and maximized at $a^* = \frac{1}{3}$. Moreover, at $a^* = 0$ and $a^* = \frac{1}{2}$, it holds that $v_1(a^*) = \pi = 2$. Thus, $v_1(a^*) > \pi$ for any $a^* \in (0, \frac{1}{2})$. However, such a contract incentivizes the agent to deviate to zero effort since this guarantees him payoff of $v_1(a^*) > \pi$ at zero costs.

Second, in the absence of moral hazard, the cheapest way to entice the agent to sign a contract that stipulates action $a$ is to offer a fixed wage of $w(a) = (\pi + a)^2$. It can be verified that profit is single-peaked in $a$ in this case. The first best action is $a^{FB} = 0.488$. Thus, the FOA suggests that the second best action is smaller than the first best action. However, the modified FOA establishes that the second best action is in reality greater than the first best action. ▲

![Figure 4: A simplified counterexample.](image)

---

17 More generally, (11) implies that $v_1(a)$ is decreasing in $a$ when $p(a)$ is concave and increasing in $a$ when $p(a)$ is convex. The opposite holds for $v_2(a)$.

\[ \text{end of the support, } a^* = 0, \text{ the principal can do no better than offering a fixed wage of } w = \pi^2 = 4. \text{ Since } p(0) = 0, \text{ the principal’s profit it then } x_1 - w = 1, \text{ coinciding with (12).} \]
Example 4 (Threshold monitoring): Continuing Example 3, assume now that the principal has access to a monitoring technology that I will term threshold monitoring. Threshold monitoring is characterized by some threshold, \( a' \), such that the agent’s action \( a \) is observable and verifiable if and only if \( a \leq a' \). The principal can thus implement actions below \( a' \) at first-best costs. Conversely, when trying to induce an action above the threshold, the principal no longer has to worry about deviations to actions below \( a' \) since such actions are now verifiable and thus punishable. Hence, there are fewer “non-local” incentive compatibility constraints. Consequently, the set of implementable actions is enlarged when \( a' \) increases.

Recall that \( p(a) = a + \frac{1}{2}(a^2 - a^3) \) is concave for all \( a \geq \frac{1}{3} \). Hence, if \( a' \geq \frac{1}{3} \) all \( a \in [a', \frac{2}{3}] \) are on the concave hull of \( p(a) \) when the support of the latter is restricted to \([a', \frac{2}{3}]\). The restriction of the support captures the fact that the only worry is to prevent deviations above the threshold. It follows that all actions above the threshold are implementable when \( a' \geq \frac{1}{3} \). The case where \( a' \in [0, \frac{1}{3}] \) is more interesting. Here, it can be shown that the concave hull of \( p(a) \) on the support \([a', \frac{2}{3}]\) consists of the interval \([\frac{1-a'}{2}, \frac{2}{3}]\). As indicated earlier, this set grows larger when \( a' \) increases. Assume the cost of the monitoring technology to the principal is \( c(a') \), with \( c'(a') > 0 \).

Threshold monitoring does not appear to have been studied before, quite possibly because it holds limited interest when the FOA is valid. Non-local incentive compatibility constraints are redundant in such a setting. Hence, there is no gain from removing these constraints. Threshold monitoring is of potential interest only because it allows low actions to be implemented at first-best costs. Since monitoring is costly, the principal pursues one of two strategies. The first strategy is to not invest in monitoring at all, \( a' = 0 \), and instead rely only on incentives to induce the target action, \( a^* \). The second strategy is to invest just enough in monitoring that the target action can be verified and thus demanded at a fixed wage, in which case \( a' = a^* \). In other words, monitoring and performance contracts will never coexist. In this sense, monitoring and incentives are perfect substitutes. Note also that if \( a' > 0 \), a decrease in the marginal costs of the technology will lead the principal to invest in (weakly) more monitoring and at the same time demand (weakly) higher effort from the agent.

The use of threshold monitoring differs when the FOA is not valid, however. Consider Figure 4 from Example 3. Without threshold monitoring, or \( a' = 0 \), the optimal action is \( a = \frac{1}{2} \). However, the principal is better off if marginally lower actions could be induced. Lower actions can be implemented if \( a' \) is increased above zero,
as explained above. On the other hand, when $a^t$ becomes large enough it is more profitable to demand action $a^t$ at a fixed wage, at least as long as $a^t$ is below the first best action, $a^{FB} = 0.488$. There is no gain to increasing $a^t$ beyond $a^{FB}$. Figure 5 describes the optimal action to induce or demand as a function of the monitoring level, $a^t$ (details are available on request).

Note that the principal uses monitoring and a performance contract at the same time if $a^t \leq 0.069$. Monitoring allows more actions to be implemented, but it is still the case that a performance contract is needed to induce the most profitable action. Note also that the principal is fully aware that the agent will never take an action below the threshold, given the accompanying contract. An outside observer would thus see the principal investing in monitoring technology, yet the technology would never catch the agent shirking. In fact, for $a^t \leq 0.069$, the induced action is $a^* = \frac{1-a^t}{2}$, which is strictly decreasing in $a^t$. Hence, as the cost of the technology decreases, the principal might invest more heavily in monitoring, yet at the same time induce a lower action. Moreover, in this particular example, any $a^t > 0$ is optimally paired with an induced action that is strictly smaller than the action that goes with $a^t = 0$. This is in part because the first best action is small compared to the second best action without monitoring; see Example 3. Stated more succinctly, the agent works less hard when he is being monitored than when he is not being monitored.

![Figure 5: Threshold monitoring.](image)

**Example 5 (Another counterexample):** Consider a setting like the one in Example 1, where $p(a)$ is convex. With a risk-neutral principal, the FOA mistakenly invokes the procedure detailed in Example 3 to derive $\pi(a^*)$ in (12). Recall that
the candidate contract is independent of \((x_1, x_2)\) for any given \(a^*\). Hence, if the good outcome is sufficiently valuable to the principal – \(x_2\) is very large – then \(\pi(a^*)\) is strictly increasing in its argument and the FOA identifies \(\pi\) as being the optimal action to implement. Under the FOA, the candidate contract satisfies the participation constraint and the local incentive compatibility constraint that \(EU''(\pi) = 0\) (see footnote 16). For instance, if, as in Araujo and Moreira’s (2001) example, \(p(a) = a^{\frac{3}{2}}, a \in [0, 0.81], v(w) = \sqrt{w}\) and \(\pi = 0\), then (11) yields the contract \((w_1, w_2) = ((\frac{27}{100})^2, (\frac{2729}{2700})^2) = (0.0729, 1.0216)\). However, the convexity of \(p(a)\) implies that the agent’s expected utility is convex in \(a\). Thus, the stationary point identified by \(EU'(\pi) = 0\) minimizes the agent’s utility. Hence, he is better off deviating to \(a\). By Proposition 3, the participation constraint must instead be paired with the incentive compatibility constraint that \(EU(\pi) \geq EU(a)\). As mentioned in Example 1, finding the optimal contract subject to these constraints is a standard problem. In Araujo and Moreira’s (2001) example the solution is \((w_1, w_2) = (0, \frac{100}{81}) = (0, 1.23457)\) rather than \((0.0729, 1.0216)\). Thus, the FOA identifies the correct optimal action when \(x_2\) is large enough, but misspecifies the optimal contract. In particular, the true optimal contract offers stronger incentives than the FOA contract in the sense that \(w_2 - w_1\) is larger than suggested by the FOA. ▲

**Example 6 (Non-Concave \(p(a)\)):** The agent produces one or more goods for the principal. The principal observes the outcome \(x\), which can be a combination of quantity and quality along a number of dimensions. The agent can use one of two machines, but he has resources to run only one them. The old machine is trivial to use, but it is less likely to produce an \(x\) that scores highly along any dimension. The agent has yet to learn how to use the new and more sophisticated machine. The agent’s action is to determine how much, if any, time and effort to invest in mastering the new machine. If he does not learn how to operate the new machine, the fallback is to run the old machine. Consider the following two thought experiments.

First, assume that the agent must in advance block off \(a\) units of time to study the machine. By blocking off this time, he forgoes opportunities even if he is lucky enough that he quickly understands how to operate the machine. The cost of \(a\) units of time is \(a\); cost are linear as in the above model. Let \(t\) denote the stochastic time at which the agent experiences the “aha moment” and masters the machine. The continuously differentiable distribution of \(t, Q(t)\), measures the probability that the agent has understood the machine by time \(t\). Let \(q(t)\) denote the density. Note that
the agent masters the machine in the allotted time if and only if \( t \leq a \), the probability of which is \( p(a) = Q(a) \). Thus, \( p'(a) = q(a) \). Consequently, \( p(a) \) is concave only if the density is decreasing. On the other hand, if \( Q(t) \) has a unimodal density, say, then \( p(a) \) is first convex and then concave in \( a \).

Second, assume instead that the agent can return to other activities as soon as he has mastered the machine. In this stopping problem, the agent has to decide on a maximum duration, \( s \), that he will devote to the new machine. In other words, \( s \) is the time at which he will stop trying. If \( t < s \), the agent incurs costs of only \( t \); he succeeds before he would have given up. Otherwise, his costs are \( s \). Expected costs are

\[
c(s) = \int_{0}^{\infty} \min\{t, s\} q(t) dt = \int_{0}^{s} t q(t) dt + s(1 - Q(s)).
\]

Note that \( c(s) \) is concave since it integrates concave functions. The intuition is that when \( s \) increases, it becomes less likely that the agent will have to wait all the way to \( s \) is reached before he understands the machine. Hence, expected costs increase at a decreasing rate, with \( c'(s) = 1 - Q(s) \geq 0 \) and \( c''(s) = -q(s) \leq 0 \). Next, normalize the problem by switching variables and thinking about the agent as deciding which expected costs to incur. Let \( a = c(s) \) denote the choice variable. With some abuse of notation, let \( s(a) \) denote the inverse of this function. Note that \( s(a) \) is increasing and convex. The probability that the agent learns how to operate the machine by incurring cost \( a \) is \( p(a) = Q(s(a)) \). By the inverse function theorem, \( p'(a) = \frac{q(s(a))}{1 - Q(s(a))} \).

Thus, \( p'(a) \) coincides with the hazard rate, i.e. the probability that the agent will understand the machine immediately, contingent on not understanding it earlier. Since \( s(a) \) is increasing, \( p(a) \) is locally concave if the hazard rate is locally decreasing and locally convex if the hazard rate is locally increasing. Note that the hazard rate may be increasing even if \( Q \) is concave. The important point here is that concavity of \( Q(t) \) does not imply concavity of \( p(a) \).\(^{18}\)

**Example 7 (Justifying the FOA):** The spanning condition has often been implicitly imposed in papers with a continuum of actions. Perhaps the most significant example of this is in LiCalzi and Spaeter (2003) who provide two classes of distributions that satisfy Rogerson’s (1985) sufficient conditions for the validity of the FOA. Thus, this paper is customarily cited in papers that rely on the FOA. The first family

\(^{18}\)For example, assume \( Q(t) = t, t \in [0, 1] \). Then, \( c(s) = \frac{1}{2} s(2 - s), s \in [0, 1] \). The cost function has range \([0, \frac{1}{2}]\) and inverse \( s(a) = 1 - \sqrt{1 - 2a}, a \in [0, \frac{1}{2}] \). Here, \( p(s(a)) = s(a) \) is convex.
of (univariate) distributions is

\[ F(x|a) = x + \tau(x)\eta(a), \quad x \in [0, 1]. \]

Obviously, conditions must be imposed on \( \tau(\cdot) \) and \( \eta(\cdot) \) to ensure that \( F(x|a) \) is a proper distribution function. As mentioned, LiCalzi and Spaeter (2003) identify additional assumptions on both \( \tau(\cdot) \) and \( \eta(\cdot) \) which ensure Rogerson’s conditions are met. Note, however, that these distribution functions are separable in \( x \) and \( a \). Thus, although it seems to not have been observed before, it should be clear that \( F(x|a) \) could be stated as in (1). Hence, the modified FOA is always valid in this family of distributions, even without LiCalzi and Spaeter’s (2003) additional assumptions.

Example 1 in Kadan and Swinkels (2013) can also be written as in (1), with \( G(x) = x, H(x) = x + x^2 - x^3, x \in [0, 1], \) and \( p(a) = 2a^2 - a^3, a \in [\frac{2}{3}, 1]. \)

Kadan and Swinkels (2013) observe that this distribution does not have the monotone likelihood ratio property. For this reason, none of the general justifications of the FOA can be invoked. However, note that \( p(a) \) is concave on \([\frac{2}{3}, 1]\). As indicated after Proposition 3, the FOA is therefore in fact valid. In this example, \( G(x) \) first order stochastically dominates \( H(x) \), or \( G(x) \leq H(x) \) for all \( x \). Jung and Kim (2015, p. 260) use an example where \( H(x) \) is a mean-preserving spread over \( G(x) \) to illustrate the fact that concavity of \( p(a) \) is sufficient for the validity of the FOA. ▲

5 Multi-dimensional effort

This section generalizes the characterization in Section 3. For any integer, \( m \geq 1 \), consider distribution functions of the form

\[ F(x|p) = \sum_{j=1}^{m} p_j G_j(x) + \left( 1 - \sum_{j=1}^{m} p_j \right) H(x), \]

where \( p \) is an element of the \( m \)-dimensional probability simplex, i.e. \( p_j \geq 0 \) for all \( j \) and \( \sum_{j=1}^{m} p_j \leq 1 \). Think of the vector \( p = (p_1, p_2, \ldots, p_m) \) as being the agent’s action. Let \( P \) denote the set of feasible \( p \) vectors. Assume \( P \) is non-empty and convex and

\[ 19 \text{Example 1 in Jewitt et al (2008) uses the Farlie-Gumbel-Morgenstern copula, } f(x|a) = 1 + \frac{1}{2}(1 - 2x)(1 - 2a), \text{ where } x, a \in [0, 1]. \text{ This distribution is also separable in } x \text{ and } a \text{ and thus can be written as in (1).} \]
let \( \text{int}(P) \) denote the interior of this set. Finally, let \( p_{-j} \) denote \( p \) with the \( j \)th element removed and write \( p = (p_j, p_{-j}) \). For all \( j \), assume that if \( (p_j, p_{-j}) \in P \) then there exists a \( p_j' \neq p_j \) such that \( (p_j', p_{-j}) \) is also in \( P \). This is a “full dimensionality” assumption, implying that no \( p_j \) is pre-determined by \( p_{-j} \). The cost function, \( c(p) \), is assumed to be continuously differentiable on \( P \). It may or may not be monotonic and no assumptions are imposed on its curvature. Hart and Holmström (1987) adopt the normalization in the current section, although they assume \( m = 1 \). Since \( F(x|p) \) is linear in \( p \), Hart and Holmström (1987) refer to (1) as the Linear Distribution Function Condition (LDFC). They also implicitly assume that \( c \) is convex. Assume that \( G_1, G_2, ..., G_m \) and \( H \) are all distinct from each other.

The agent’s expected utility can be written

\[
\begin{align*}
EU(p) &= \sum_{j=1}^{m} p_j \beta_j + \gamma - c(p),
\end{align*}
\]

where

\[
\begin{align*}
\beta_j &= \int v(w(x))d(G_j(x) - H(x)), \quad \gamma = \int v(w(x))dH(x).
\end{align*}
\]

For any interior \( p^* \in \text{int}(P) \), L-IC\(_p^* \) implies that

\[
\beta_j = c_j(p^*) \text{ for all } j = 1, 2, ..., m,
\]

where \( c_j(p) \) denotes the partial derivative of \( c(p) \) with respect to \( p_j \). Note that it may be impossible to satisfy (15). Imagine for example that there are a finite number, \( k \), of possible outcomes. Thus, a contract specifies \( k \) wages. By (14), each \( \beta_j \) is essentially a weighted average of these. Thus, if \( k \) is strictly lower than \( m \) the principal will typically not have enough degrees of freedom to satisfy \( \beta_j = c_j(p^*) \) for all \( j \).

In the following I therefore restrict attention to those \( p \in P \) for which a L-IC\(_p \) contract exists. Let \( P^L \) denote this set. In the multi-dimensional case, this set may depend on \( G_1, G_2, ..., G_m \) and \( H \). For simplicity, I focus on \( p \in \text{int}(P) \) as well.

Thus, assume \( p^* \in P^L \cap \text{int}(P) \). Then, utilizing (15),

\[
\begin{align*}
EU(p^*) - EU(p) &= c(p) - \left[ c(p^*) + \sum_{j=1}^{m} (p_j - p_j^*) c_j(p^*) \right]
\end{align*}
\]

for all \( p \in P \). The term in square brackets is the tangent plane to \( c(p) \) through the
point $p^*$. Let $P^C$ denote the subset of $P$ for which $c(p)$ coincides with its convex hull. Note that $P^C$ does not depend on $G_1, G_2, ..., G_m$ and $H$. Given the change in normalization, $P^C$ is the natural counterpart to $A^C$ in Section 3. By definition of the convex hull, $p^* \in P^C$ if and only if (16) is non-negative for any $p$.

**Proposition 4** Assume $p^* \in \text{int}(P)$. Then, there exists a G-IC$_{p^*}$ contract (that yields bounded utility) if and only if $p^* \in P^L \cap P^C$.

**Proof.** If $p^* \in P^L$, then (16) must hold. Moreover, if $p^* \in P^C$, then there is no profitable deviation since (16) is always non-negative. Thus, the contract is G-IC$_{p^*}$. Conversely, if $p^* \notin P^L \cap P^C$ then either there is no L-IC$_{p^*}$ contract or there is an incentive to deviate from the L-IC$_{p^*}$ contract. In either case, the contract does not satisfy G-IC$_{p^*}$. ■

As in the one-dimensional case, (16) immediately implies that L-IC$_{p^*}$ is sufficient for G-IC$_{p^*}$ on $P^C$ (however, a L-IC$_{p^*}$ contract exists only on the subset $P^L \cap P^C$).

**Proposition 5** If $p^* \in P^C$ then any L-IC$_{p^*}$ contract is G-IC$_{p^*}$.

Once again, a modified FOA suggests itself. First, Proposition 4 is used to identify the (interior of the) feasible set. Then, the FOA is applied to this set. By Proposition 5, the solution, if one exists, is G-IC. This result now establishes a counterpart to results in the standard FOA literature. There, it is often assumed that the second best action is interior, and sufficient conditions are then provided under which the solution to the FOA program identifies this second best action.

**Proposition 6** Assume the second best action, $p^*$, is in $\text{int}(P)$. Then, the modified FOA is valid. That is, applying the FOA to $P^L \cap P^C$ yields $p^*$ and characterizes the optimal contract.

The model reduces to one in which there are $m + 1$ outcomes if $G_1, G_2, ..., G_m$, and $H$ are degenerate at different points. In this case, a contract consists of $m + 1$ wages. At the same time, the $m$ constraints in (15) along with the participation constraint yield $m + 1$ conditions. Hence, the optimal contract can be derived from the constraints alone. Examples 1-5 in Section 4 are special cases, with two outcomes
and \( m = 1 \). Example 3 solves for the optimal contract in that case. In similar fashion, when \( m \geq 1 \) and \( p^* \in P^L \cap P^C \), the \( m + 1 \) constraints yield

\[
\begin{align*}
v_i &= \bar{u} + c(p^*) + (1 - p^*_i) c_i(p^*) - \sum_{j \neq i} p^*_j c_j(p^*), \quad i = 1, 2, ..., m \\
v_{m+1} &= \bar{u} + c(p^*) - \sum_{j=1}^{m} p^*_j c_j(p^*),
\end{align*}
\]

where \( v_i \) and \( v_{m+1} \) are the utilities achieved if the outcome is drawn from \( G_i \) and \( H \), respectively. Wages can be obtained by inverting \( v(w) \). Since wages are determined only by the constraints, there is a unique contract that implements \( p^* \) with a binding participation constraint.

6 Conclusion

Although the spanning condition appears simple and dates back to Grossman and Hart (1983), the first full characterization of its solution is given here. The characterization is surprisingly simple. Nevertheless, as the examples in Section 4 demonstrate, the results enable substantial economic insights. For instance, Mirrlees’ (1999) famous counterexample that the FOA is not always valid can be solved using the techniques presented here. Indeed, the paper’s insights make it possible to construct a much simpler counterexample.

The literature relies heavily on the FOA for applications that assume a continuum of actions. One reason is that until now no tractable workhorse model has been identified that enables one to get a handle on the general incentive compatibility problem. One key advantage of the spanning condition is that it yields a tractable model in which it is easy to identify precisely which non-local incentive compatibility constraints are binding. This level of detail may make the model ideal as a laboratory of sorts in which to examine comparative statics when the FOA is not valid. As an example, a very stark type of monitoring is used to demonstrate that comparative statics may be markedly different when the FOA is not valid.

The characterization can be extended to a more general version of the spanning condition in which the agent determines the weights placed on several technologies. Given a “full dimensionality” assumption, the modified FOA remains valid.
References


29


