THE GENERALIZED INFORMATIVENESS PRINCIPLE

Pierre Chaigneau
Alex Edmans
Daniel Gottlieb

The Generalized Informativeness Principle
Pierre Chaigneau, Alex Edmans, and Daniel Gottlieb
NBER Working Paper No. 20729
December 2014
JEL No. D86,J33

ABSTRACT

This paper shows that the informativeness principle, as originally formulated by Holmstrom (1979), does not hold if the first-order approach is invalid. We introduce a "generalized informativeness principle" that takes into account non-local incentive constraints and holds generically, even without the first-order approach. Our result holds for both separable and non-separable utility functions.

Pierre Chaigneau
HEC Montreal
3000 chemin de la côte-Sainte-Catherine
Montreal H3T 2A7
Canada
pierre.chaigneau@hec.ca

Alex Edmans
The Wharton School
University of Pennsylvania
2460 Steinberg Hall - Dietrich Hall
3620 Locust Walk
Philadelphia, PA 19104
and NBER
aedmans@wharton.upenn.edu

Daniel Gottlieb
The Wharton School
University of Pennsylvania
3303 Steinberg Hall-Dietrich Hall
3620 Locust Walk
Philadelphia, PA 19104
dgott@wharton.upenn.edu

1 Introduction

The informativeness principle, or sufficient statistic theorem, states that a signal has positive value if and only if it affects the likelihood ratio. This principle is believed to be the most robust result from the moral hazard literature. For example, the textbook of Bolton and Dewatripont (2005) states that this literature has produced very few general results, but the informativeness principle is one of the few results that is general.[1]

Due to its perceived robustness, the informativeness principle has been applied to many settings. It is the key concept behind the theories of relative performance evaluation (Baiman and Demski (1980), Holmstrom (1982)), tournaments (Lazear and Rosen (1981)), and yardstick competition (Shleifer (1985)). This wide applicability in turn has led to substantial impact in many fields, such as compensation, insurance, and regulation. For example, in Bebchuk and Fried's (2004) influential book on the inefficiency of executive compensation practices, one of their leading arguments is that exogenous "luck" is not filtered out from the contract.

The original formulation of the informativeness principle, in Holmstrom (1979) and Shavell (1979), assumes the validity of the first-order approach ("FOA"): that the agent's incentive constraint can be replaced by its first-order condition. All of its generalizations assume either the FOA (e.g. Gjesdal (1982), Amershi and Hughes (1989), Kim (1995)) or that the agent chooses between two actions only (e.g. Hart and Holmstrom (1987), Bolton and Dewatripont (2005)). As is well-known, the FOA is generally not valid.[2] Assuming only two actions has a similar effect to using the FOA, as it means that only one incentive constraint binds, but is unrealistic.

Due to the significance of the informativeness principle and the restrictive setting in which it was derived, it is important to understand whether it is a robust property that holds more generally. In this paper, we show that the informativeness principle may not hold when the FOA is invalid. Our main contribution is to propose a "generalized

---

[1] They write: "The basic moral hazard problem has a fairly simple structure, yet general conclusions have been difficult to obtain ... Among the main general predictions of the model is the informativeness principle" (p129 and p169).

[2] Rogerson (1985) derives the most well-known sufficient conditions for the validity of the FOA in the single-signal case. As Jewitt (1988) points out, these assumptions are so strong that they are not satisfied by any standard distribution. Moreover, they are no longer sufficient if the principal observes multiple signals, which is needed to analyze the informativeness principle (as the principal observes output and an additional signal). Jewitt (1988), Sinclair-Desgagné (1994), Conlon (2008), and Ke (2012) obtain sufficient conditions for the validity of the FOA in the mutliple-signal case.

informativeness principle" that provides sufficient conditions for a signal to have value and holds generically (i.e., for all parameters except for sets of measure zero).

Since the original informativeness principle assumed the FOA, only the likelihood ratio between adjacent efforts mattered. When the FOA is invalid, the binding incentive constraint(s) are not local. Thus, even if the signal is informative about the local likelihood ratio, it may have zero value for the contract. Since we do not know from the outset which incentive constraints will bind, our generalized informativeness principle requires the signal to affect the likelihood ratio between the principal's preferred effort and all other efforts, rather than only adjacent efforts.

When only one incentive constraint binds, the generalized informativeness principle always holds. If the signal affects all incentive constraints, it will affect the binding incentive constraint, and thus have strictly positive value by the same intuition as in Holmstrom (1979). The principal can use the signal to relax the binding incentive constraint by transferring payments from states with low likelihood ratios to states with high likelihood ratios, in turn reducing the expected payment.

With more than two efforts, however, multiple incentive constraints bind for an open set of parameters. In this case, signals that affect all likelihood ratios may still have zero value. While the principal can use such a signal to transfer payments to relax one binding constraint, the same transfer may tighten another binding constraint by the same magnitude, and so the overall payment reduction is exactly zero.

Counter-examples such as this are knife-edge in that they require the benefit from relaxing one binding constraint to exactly equal the cost of tightening other binding constraints. Intuitively, they require the shadow prices of the binding constraints to be equal, and so they are non-robust to small perturbations in the probability distribution or the utility function. Accordingly, we show that, except for a set of parameters with measure zero, any signal that affects all likelihood ratios has positive value.

This generalized informativeness principle allows not only for additively separable utility (as in all previous versions of the principle), but also for multiplicatively separable and non-separable utility. This generalization, however, requires us to distinguish between efforts that can be implemented with no agency costs (i.e. a constant wage, so that the first best can be achieved) and those that cannot. Previous versions of the informativeness principle assume additively separable utility (in which case only the cheapest effort can be implemented with a constant wage) and an interior effort (so that it cannot be implemented with a constant wage). With non-separable utility,

many efforts may be implementable with a constant wage, in which case the signal automatically has no value. We show that, for both separable and non-separable utility, the informativeness principle (generically) holds only for efforts that cannot be implemented with a constant wage.

Finally, Holmstrom's (1979) original theorem was an "if and only if" result, providing necessary and sufficient conditions for a signal to have strictly positive value in contracting. While our main result concerns the more surprising ("sufficient") part, we also generalize the "necessary" part of his theorem. That is, we show that an uninformative signal has no value for the contract even when utility is non-separable.

Our paper proceeds as follows. Section 2 revisits the original informativeness principle and shows that it may not hold if the FOA is invalid. Section 3 shows that a generalized informativeness principle generically holds. Section 4 concludes.

2   The Informativeness Principle and the FOA

This section shows that the informativeness principle needs to be modified when non-local incentive constraints bind. There is a single risk-neutral principal ("she") and a single risk-averse agent ("he"). The agent chooses an action $e \in \mathcal{E}$, which we refer to as "effort" and is not observable by the principal. The principal observes output $x \in \mathcal{X}$ and a signal $s \in \mathcal{S}$, which may be informative about e. Both output and the signal are contractible. We refer to a pair (x, s) as a "state." In Section 3, we will assume that the action, output, and signal spaces are finite; for now, to achieve comparability with Holmstrom (1979), we allow them to be intervals of the real line as well.

While Holmstrom (1979) considers an additively separable utility function, we follow Grossman and Hart (1983) and generalize to the following utility function:

Assumption 1. *The agent's Bernoulli utility function over income w and effort e is*

$$U(w, e) = G(e) + K(e) V(w). \tag{1}$$

(i) $K(e) > 0$ *for all e;* **(ii)** $V : \mathcal{W} \to \mathbb{R}$ *is continuously differentiable, strictly increasing, and strictly concave, and $\mathcal{W} = (\underline{w}, +\infty)$ is an open interval of the real line (possibly with $\underline{w} = -\infty$); and* **(iii)** $U(w_1, e_1) \geq U(w_1, e_2) \implies U(w_2, e_1) \geq U(w_2, e_2)$ *for all $e_1, e_2 \in \mathcal{E}$ and $w_1, w_2 \in \mathcal{W}$.*

4

The agent has utility function (1) if and only if his preferences over income lotteries are independent of his effort. Conditions (i) and (ii) state that the agent likes money and dislikes risk. Condition (iii) requires preferences over known efforts to be independent of income.[3] When $K(e) = \bar{K}$ for all e, the utility function is additively separable between effort and income as in Holmstrom (1979). When $G(e) = 0$ for all e, it is multiplicatively separable.[4] The agent's reservation utility is $\bar{U}$.

As Grossman and Hart (1983) show, the principal's problem can be split in two stages. First, she finds the cheapest contract that induces each effort $e \in \mathcal{E}$. Second, she determines which effort e to induce. This paper focuses on the first stage: whether the principal can use the signal s to reduce the cost of implementing a given effort.[5]

First, we state Holmstrom's (1979) original theorem[6]:

Theorem. (Informativeness Principle): *Assume that the utility function is additively separable and that the FOA is valid. Suppose states are distributed according to a continuously differentiable probability density function $f(x, s|e)$. The signal has zero value for implementing e^* if and only if*

$$\frac{f_e(x, s|e^*)}{f(x, s|e^*)} = \phi(x, e^*) \tag{2}$$

for almost all x, s.

The expression on the right of (2) corresponds to the change in the likelihood ratio $\frac{f(x,s|e^*+\Delta e)}{f(x,s|e^*)}$ for infinitessimal changes in effort $\Delta e \approx 0$. Since only the local IC matters when the FOA is valid, the value of the signal only depends on this local effect.

We now present an example in which the signal is informative (i.e., (2) fails to hold) and yet has zero value because the FOA is not valid and so the relevant IC is not local. This violation motivates the generalized informativeness principle of Section 3.[7]

---

[3]Assumption 1(iii) still allows the agent's preferences for *lotteries* over effort to depend on income.

[4]Multiplicative separability is commonly used in macroeconomics (e.g. Cooley and Prescott (1995)). In addition, Edmans, Gabaix, and Landier (2009) show that they are necessary and sufficient to obtain empirically consistent scalings of CEO incentives with firm size.

[5]Holmstrom (1979) avoids this issue by assuming that either the signal is informative for all effort levels or for no effort level.

[6]In Supplementary Appendix B.2 we formally define what it means for a signal to have value.

[7]We are not the first to point out that signals that only affect non-binding ICs have zero value (see, e.g., footnote 7 in Holmstrom and Milgrom (1987)). We are, however, the first to show that even signals that affect all ICs (including the binding ones) may have zero value: see Example 2.

Example 1. *Assume that the utility function is additively separable between income and effort: $U(w, e) = V(w) + G(e)$. The utility of income V is bounded above by $\widehat{U}$ and the cost of effort is*

$$G(e) = \begin{cases} 0 & \text{if } e = 0 \\ -K & \text{if } e \notin \{0, 1\} \\ -C & \text{if } e = 1 \end{cases},$$

where $K \geq \widehat{U}$ and $C > 0$.

Since the cost of any effort $e \notin \{0, 1\}$ exceeds the maximum utility of income $\widehat{U}$, the agent will never choose $e \notin \{0, 1\}$. If the principal wishes to implement $e = 1$, the relevant IC is the one preventing the agent from selecting $e = 0$. The signal has zero value for implementing $e = 1$ if and only if the likelihood ratio between efforts 0 and 1 is independent of the signal, i.e.:

$$\frac{f(x, s|1)}{f(x, s|0)} = \frac{f(x, s'|1)}{f(x, s'|0)}$$

for almost all (x, s) and (x, s'). This condition does not imply and is not implied by the local condition $\frac{f_e(x, s|1)}{f(x, s|1)} = 0$.

When the FOA is not valid, non-local ICs may bind. Then, as Example 1 shows, the relevant likelihood ratio is the one comparing the implemented effort ($e = 1$) to the effort exerted in the binding IC ($e = 0$). Thus, affecting the likelihood ratio for adjacent efforts is no longer sufficient for a signal to have positive value.

Note that Holmstrom's (1979) informativeness principle is an "if and only if" result. The less surprising part shows that uninformative signals have zero value ("necessity"). The more interesting part shows that every informative signal has strictly positive value ("sufficiency"). The main contribution of this paper is to generalize the sufficiency part. However, before doing so, we first generalize the necessity part to settings in which the FOA is not valid and the utility function is not additively separable. The proof is in Supplementary Appendix B.4.

Proposition 1. *Let (x, s) be either continuously or discretely distributed, and let $f(x, s|e)$ denote either the probability density function or the probability mass function. Suppose $\frac{f(x, s|e)}{f(x, s|e^*)} = \phi_{e^*}(x, e)$ for all e and almost all (x, s) under e^*. Then, the signal has zero value in implementing e^*.*

Thus, uninformative signals have zero value even when utility is not additively separable or the FOA is not valid. The remainder of our paper focuses on the sufficiency part.

3 The Generalized Informativeness Principle

Following Grossman and Hart (1983), we assume that states and efforts are finite: $\mathcal{E} \equiv \{1, \ldots, E\}$, $\mathcal{X} \equiv \{x_1, \ldots, x_X\}$, and $\mathcal{S} \equiv \{1, \ldots, S\}$.[8] The probability of observing state (x, s) conditional on effort e is denoted by $p_{x,s}^e \equiv \Pr(\tilde{x} = x, \tilde{s} = s | \tilde{e} = e) > 0$.

Let $h \equiv V^{-1}$ denote the inverse utility function. Since u is increasing and strictly concave, h is increasing and strictly convex. Letting $u_{x,s} \equiv u(w_{x,s})$, the principal's program can be written in terms of "utils":

$$\min_{\{u_{x,s}\}} \sum_{x,s} p_{x,s}^{e^*} h(u_{x,s}) \tag{3}$$

subject to

$$G(e^*) + K(e^*) \sum_{x,s} p_{x,s}^{e^*} u_{x,s} \geq \bar{U}, \tag{4}$$

$$\sum_{x,s} \left(K(e^*) p_{x,s}^{e^*} - K(e) p_{x,s}^e \right) u_{x,s} \geq G(e) - G(e^*) \quad \forall e \tag{5}$$

where (4) and (5) are the agent's participation and incentive constraints (IR and IC).

We first note that a signal can only have positive value when there are agency costs. Let $\bar{w}_e$ denote the wage that gives the agent his reservation utility if he exerts effort e:

$$\bar{w}_e = h\left(\frac{\bar{U} - G(e)}{K(e)} \right).$$

The principal can implement effort e^* with no agency costs if all ICs are satisfied when she offers the constant wage $\bar{w}_{e^*}$ that satisfies the IR with equality:

$$U(\bar{w}_{e^*}, e^*) \geq U(\bar{w}_{e^*}, e) \quad \forall e. \tag{6}$$

We say that the first best is feasible for e^* if condition (6) holds. The principal then obtains the first-best payoff by using a constant wage and so signals automatically have zero value. When utility is either additively or multiplicatively separable, the first best is only feasible for the least costly effort. With non-separable utility, however, it may be feasible for several different efforts. (The first-best is never achieved in Holmstrom (1979) because he assumes additively separable utility and an interior e.) The informativeness principle does not hold if the first-best is feasible, i.e. no IC binds. Remark 1 notes that it holds whenever exactly one IC binds:

---

[8]Finite efforts allow us to use Kuhn-Tucker methods to obtain necessary optimality conditions.

Remark 1. *Suppose that exactly one IC binds in Program (3)-(5) and let e^* be an effort for which the first best is not feasible. The necessary Kuhn-Tucker conditions from the principal's program yield, for all (x, s) in the support,*

$$-h'\left(u_{x,s}\right) + \mu\left(K\left(e^*\right) - K(e')\frac{p_{x,s}^{e'}}{p_{x,s}^{e^*}}\right) + \lambda K\left(e^*\right) = 0, \tag{7}$$

where $\mu \geq 0$ is the multiplier associated with the binding IC. Subtracting these conditions in states (x, s) and (x, s') gives

$$h'\left(u_{x,s}\right) - h'\left(u_{x,s'}\right) = \mu K(e')\left(\frac{p_{x,s'}^{e'}}{p_{x,s'}^{e^*}} - \frac{p_{x,s}^{e'}}{p_{x,s}^{e^*}}\right). \tag{8}$$

If $\mu = 0$, then (8) implies a constant wage, which contradicts our assumption that the first best is not feasible.[9] Therefore, $\mu > 0$ and, because $K\left(e\right) > 0$ for all e, it follows from (8) and the convexity of h that $u_{x,s} \neq u_{x,s'}$ whenever $\frac{p_{x,s'}^{e'}}{p_{x,s'}^{e}} \neq \frac{p_{x,s}^{e'}}{p_{x,s}^{e}}$.

The final case to consider is when multiple ICs bind. When there are at least three states, it is not unusual for multiple ICs to bind. Formally, we show in Supplementary Appendix B.3 that multiple ICs bind for a non-empty and open set of parameter values. Since any non-trivial model with informative signals requires at least three states (at least two outputs and at least two signals conditional on at least one output), it is important to study the case of multiple binding ICs.

We start with an example showing that, if multiple ICs bind, the generalized informativeness principle may not hold. Notice that our example follows Holmstrom (1979) and the subsequent literature in assuming additive separability:

Example 2. *There are three efforts, two outputs, and two signals: $\mathcal{E} = \{1, 2, 3\}$, $\mathcal{X} = \{0, 1\}$, and $\mathcal{S} = \{0, 1\}$. Let $K\left(1\right) = K\left(2\right) = K\left(3\right) = 1$, $G\left(1\right) = G\left(2\right) = 0$, and $G\left(3\right) = -1$. Thus, $e = 1$ and $e = 2$ both cost zero and $e = 3$ costs one. The reservation utility is $\overline{U} = 0$.*

Conditional on $e = 3$, states are uniformly distributed: $p_{x,s}^3 = \frac{1}{4} \; \forall \; x, s$. For $e \in \{1, 2\}$, the conditional probabilities are:

$$p_{1,0}^1 = p_{1,1}^2 = \frac{1}{4}, \quad p_{1,1}^1 = p_{1,0}^2 = \frac{1}{8}, \quad p_{0,0}^1 = p_{0,1}^1 = p_{0,0}^2 = p_{0,1}^2 = \frac{5}{16}.$$

---

[9] Since the agent's preferences over efforts are independent of income (Assumption (1iii)), effort e^* can be implemented with the minimum constant wage $\bar{w}_{e^*}$ if and only if it can be implemented with any other wage $w \geq \bar{w}_{e^*}$.

Note that the likelihood ratios between any two efforts are not constant:

$$\frac{p_{1,1}^3}{p_{1,1}^2} = 1 \neq 2 = \frac{p_{1,0}^3}{p_{1,0}^2}, \quad \frac{p_{1,1}^3}{p_{1,1}^1} = 2 \neq 1 = \frac{p_{1,0}^3}{p_{1,0}^1}, \quad \frac{p_{1,1}^2}{p_{1,1}^1} = 2 \neq \frac{1}{2} = \frac{p_{1,0}^2}{p_{1,0}^1}.$$

Let $e = 3$ be the effort to be implemented. The principal's program is

$$\min_{\{u_{x,s}\}} h(u_{1,0}) + h(u_{1,1}) + h(u_{0,0}) + h(u_{0,1})$$

subject to the IR and the two ICs:

$$\frac{u_{1,0} + u_{1,1} + u_{0,0} + u_{0,1}}{4} - 1 \geq 0$$

$$\frac{u_{1,0} + u_{1,1} + u_{0,0} + u_{0,1}}{4} - 1 \geq \frac{u_{1,0}}{4} + \frac{u_{1,1}}{8} + \frac{5}{16}(u_{0,0} + u_{0,1})$$

$$\frac{u_{1,0} + u_{1,1} + u_{0,0} + u_{0,1}}{4} - 1 \geq \frac{u_{1,1}}{4} + \frac{u_{1,0}}{8} + \frac{5}{16}(u_{0,0} + u_{0,1}).$$

Rewrite the constraints as

$$u_{1,0} + u_{1,1} + u_{0,0} + u_{0,1} \geq 4 \tag{9}$$

$$2u_{1,1} - (u_{0,0} + u_{0,1}) \geq 16 \tag{10}$$

$$2u_{1,0} - (u_{0,0} + u_{0,1}) \geq 16. \tag{11}$$

The solution must entail $u_{0,0} = u_{0,1}$ since, if they were different, replacing them both by $\frac{u_{0,0} + u_{0,1}}{2}$ would keep all constraints unchanged and reduce the objective function (by convexity of h). The solution must also entail $u_{1,0} = u_{1,1}$. To see this, let $(u_{0,0}, u_{0,1}, u_{1,0}, u_{1,1})$ be a solution and consider the vector that replaces $u_{1,0}$ and $u_{1,1}$ by their average $\frac{u_{1,0} + u_{1,1}}{2}$. Since the original vector was a solution, it satisfied the IR (equation (9)) and both ICs (equations (10) and (11)). Since the new vector gives the same expected utility, it also satisfies the IR (9). Moreover, taking the average between (10) and (11) establishes that the new vector is also incentive compatible:

$$u_{1,0} + u_{1,1} - (u_{0,0} + u_{0,1}) \geq 16.$$

Thus, even though the likelihood ratio is not constant for all efforts, the signal has zero value: $u_{x,0} = u_{x,1}$ for $x \in \{0,1\}$.

The intuition for the failure of the informativeness principle is as follows. For $e = 2$, the likelihood ratio at state $(1,0)$ is twice as large as at $(1,1)$. To relax the second IC

(11), we should increase $u_{1,0}$ and decrease $u_{1,1}$. For $e = 1$, the likelihood ratio at state $(1, 1)$ is twice as large as at $(1, 0)$. To relax the first IC (10), we should increase $u_{1,1}$ and decrease $u_{1,0}$. Since both the likelihood ratios $\frac{p_{1,0}^3}{p_{1,0}^2}$ and $\frac{p_{1,1}^3}{p_{1,1}^1}$ and the costs of efforts 1 and 3 coincide, the shadow prices of both ICs are the same. Thus, the benefit from relaxing one IC is exactly the same as the cost from tightening the other one. As a result, it is optimal not to make the agent's utility depend on the signal.

This result requires that the shadow prices of the binding ICs exactly coincide. If we perturb either the probabilities or the utility function slightly, the benefit from relaxing each constraint will differ. We can then improve the contract by increasing utility in the state with the highest likelihood ratio under the effort associated with the IC with the highest shadow cost. This intuition suggests that counterexamples such as the one in Example 2 are non-generic. We now establish that this is indeed the case.

To establish results that can be applied to settings with additive and multiplicative separability, we hold either K or G fixed in our economy parametrization. Therefore, we refer to an economy as either a vector of parameters $(K(e), p_{s,x}^e)_{s,x,e}$ (which holds $G(e)$ fixed), or a vector of parameters $(G(e), p_{s,x}^e)_{s,x,e}$ (which holds $K(e)$ fixed). Our results still hold if we parametrize an economy by K, G, and p. However, in this case, economies with additive or multiplicative separability are non-generic.

Theorem 1 is the main result of our paper. It states that, generically, signals that are informative about deviations to all efforts have positive value:

Theorem 1. (Generalized Informativeness Principle) *Let e^* be an effort for which the first best is not feasible. For all economies except for a set of Lebesgue measure zero, if $\frac{p_{x,s}^e}{p_{x,s}^{e^*}} \neq \frac{p_{x,s'}^e}{p_{x,s'}^{e^*}}$ for all e, then the signal has positive value.*

4   Conclusion

This paper shows that the informativeness principle may not hold when the first-order approach is violated. We establish a generalization that gives sufficient conditions for a signal to have positive value and is generically true. Our generalized informativeness principle requires the signal to affect the likelihood ratio between the implemented effort and all other efforts, and that the effort cannot be implemented with a constant payment. Our results hold for both separable and non-separable utility functions.

References

[1] Amershi, Amin H. and John S. Hughes (1989): "Multiple signals, statistical sufficiency, and Pareto orderings of best agency contracts." *RAND Journal of Economics* 20, 102–112.

[2] Baiman, Stanley, and Joel S. Demski (1980): "Economically optimal performance evaluation and control systems." *Journal of Accounting Research* 18, 184–220.

[3] Bebchuk, Lucian Arye, and Jesse M. Fried (2004): *Pay Without Performance: The Unfulfilled Promise of Executive Compensation* (Harvard University Press, Cambridge).

[4] Bolton, Patrick, and Mathias Dewatripont (2005): *Contract Theory* (MIT Press, Cambridge).

[5] Conlon, John R. (2009): "Two new conditions supporting the first-order approach to multisignal princpal-agent problems." *Econometrica* 77, 249–278.

[6] Cooley, Thomas and Edward C. Prescott (1995): "Economic growth and business cycles" in Thomas Cooley (ed.) *Frontiers in Business Cycle Research* (Princeton University Press, Princeton).

[7] Edmans, Alex, Xavier Gabaix, and Augustin Landier (2009): "A multiplicative model of optimal CEO incentives in market equilibrium." *Review of Financial Studies* 22, 4881–4917.

[8] Gjesdal, Frøystein (1982): "Information and incentives: the agency information problem." *Review of Economic Studies* 49, 373–390.

[9] Grossman, Sanford J., and Oliver D. Hart (1983): "An analysis of the principal-agent problem." *Econometrica* 51, 7–45.

[10] Hart, Oliver and Bengt Holmstrom (1987): "The theory of contracts" in Truman F. Bewley (ed.) *Advances in Economic Theory* (Cambridge University Press, Cambridge).

[11] Holmstrom, Bengt (1979): "Moral hazard and observability." *Bell Journal of Economics* 10, 74–91.

[12] Holmstrom, Bengt (1982): "Moral hazard in teams." *Bell Journal of Economics* 13, 326–340.

[13] Holmstrom, Bengt and Paul R. Milgrom (1987): "Aggregation and linearity in the provision of intertemporal incentives." *Econometrica* 55, 303–328.

[14] Jewitt, Ian (1988): "Justifying the first-order approach to principal-agent problems." *Econometrica* 56, 1177–1190.

[15] Ke, Rongzhu (2012): "A fixed-point method for validating the first-order approach." Working paper, Chinese University of Hong Kong.

[16] Kim, Son Ku (1995): "Efficiency of an information system in an agency model." *Econometrica* 63, 89–102.

[17] Lazear, Edward P. and Sherwin Rosen (1981): "Rank-order tournaments as optimum labor contracts." *Journal of Political Economy* 89, 841–864.

[18] Rogerson, William P. (1988): "The first-order approach to principal-agent problems." *Econometrica* 53, 1357–1368.

[19] Shavell, Steven (1979): "Risk sharing and incentives in the principal and agent relationship." *Bell Journal of Economics* 10, 55–73.

[20] Shleifer, Andrei (1985): "A theory of yardstick competition." *RAND Journal of Economics* 16, 319–327.

[21] Sinclair-Desgagné, Bernard (1994): "The first-order approach to multi-signal principal-agent problems." *Econometrica* 62, 459–465.

A   Proof of Theorem 1

Throughout the proof, we will use bold letters to denote vectors. We will use the following corollary of Sard's Theorem:

Corollary 1. (Sard) *Let $X \subset \mathbb{R}^n$ and $\Theta \subset \mathbb{R}^p$ be open, $F : X \times \Theta \to \mathbb{R}^m$ be continuously differentiable, and let $n < m$. Suppose that for all (x, θ) such that $F(x, \theta) = 0$, $DF(x, \theta)$ has rank m. Then, for all θ except for a set of Lebesgue measure zero, $F(x, \theta) = 0$ has no solution.*

For simplicity, suppose that only two ICs bind; it is straightforward but notationally cumbersome to generalize the analysis for more than two binding ICs. Without loss of generality (renumbering efforts if necessary), let $e^* = 3$ denote the implemented effort, and let $e = 1$ and $e = 2$ denote the two efforts with binding ICs. By assumption, the first best is not feasible for $e^* = 3$. The principal's program is

$$\min_{u_{x,s}} \sum_{x=x_1}^{x_X} \sum_{s=1}^{S} p_{x,s}^{e^*} h(u_{x,s}) \tag{12}$$

$$\text{subject to} \quad G(e^*) + K(e^*) \sum_{x=x_1}^{x_X} \sum_{s=1}^{S} p_{x,s}^{e^*} u_{x,s} \geq \bar{U}, \tag{13}$$

$$G(e^*) + K(e^*) \sum_{x=x_1}^{x_X} \sum_{s=1}^{S} p_{x,s}^{e^*} u_{x,s} \geq G(e) + K(e) \sum_{x=x_1}^{x_X} \sum_{s=1}^{S} p_{x,s}^{e} u_{x,s} \quad \forall e \tag{14}$$

There are two possible cases depending on whether the IR (13) binds. Here, we consider the case where it binds. The case where it does not bind is analogous and is presented in Supplementary Appendix B.4.

The (necessary) first-order condition with respect to $u_{x,s}$ is

$$-p_{x,s}^{e^*} h'(u_{x,s}) - \mu_1 K(1) p_{x,s}^1 - \mu_2 K(2) p_{x,s}^2 + \lambda K(e^*) p_{x,s}^{e^*} = 0 \quad \forall x, s. \tag{15}$$

Following the parametrization of an economy, we keep either $\mathbf{G} \equiv (G(3), G(2), G(1))$ or $\mathbf{K} \equiv (K(3), K(2), K(1))$ constant. Accordingly, let either $\Theta = \mathbf{K}$ (if $\mathbf{G}$ is being held constant) or $\Theta = \mathbf{G}$ (if $\mathbf{K}$ is being held constant).

For the agent's payments to be independent of the signal, the system of equations (13), (14), and (15) must have $u_{x,s} = u_x$ as a solution for all x, s. Combining these equations, they can be written as $F(\mathbf{u}, \mu_1, \mu_2, \lambda_3; \Theta, \mathbf{p}) = 0$, where

$$
F\left(\underbrace{\mathbf{u}}_{X},\underbrace{\mu_1,\mu_2,\lambda}_{3};\underbrace{\Theta}_{3},\underbrace{\mathbf{p}}_{3XS}\right) \equiv
\begin{bmatrix}
p_{1,1}^3 h'(u_1) + \mu_1 K(1)p_{1,1}^1 + \mu_2 K(2)p_{1,1}^2 - \lambda K(3)p_{1,1}^3 \\
p_{1,2}^3 h'(u_1) + \mu_1 K(1)p_{1,2}^1 + \mu_2 K(2)p_{1,2}^2 - \lambda K(3)p_{1,2}^3 \\
\vdots \\
p_{1,S}^3 h'(u_1) + \mu_1 K(1)p_{1,S}^1 + \mu_2 K(2)p_{1,S}^2 - \lambda K(3)p_{1,S}^3 \\
\vdots \\
p_{X,1}^3 h'(u_X) + \mu_1 K(1)p_{X,1}^1 + \mu_2 K(2)p_{X,1}^2 - \lambda K(3)p_{X,1}^3 \\
p_{X,2}^3 h'(u_X) + \mu_1 K(1)p_{X,2}^1 + \mu_2 K(2)p_{X,2}^2 - \lambda K(3)p_{X,2}^3 \\
\vdots \\
p_{X,S}^3 h'(u_X) + \mu_1 K(1)p_{X,S}^1 + \mu_2 K(2)p_{X,S}^2 - \lambda K(3)p_{X,S}^3 \\
\sum_{x=1}^{X} u_x K(3) \sum_s p_{x,s}^3 + G(3) - \bar{U} \\
\sum_{x=1}^{X} u_x K(2) \sum_s p_{x,s}^2 + G(2) - \bar{U} \\
\sum_{x=1}^{X} u_x K(1) \sum_s p_{x,s}^1 + G(1) - \bar{U}
\end{bmatrix},
$$

and the terms under brackets indicate the number of elements. The remainder of the proof verifies that DF has full row rank so we can apply Corollary 1.

Write the derivative of F as:

$$
DF = \begin{bmatrix}
A & C & D_\Theta & H_3 & H_2 & H_1 \\
B & \mathbf{0_{3\times 3}} & E_\Theta & J_3 & J_2 & J_1
\end{bmatrix},
$$

where the terms inside the matrix will be defined below.

The $XS \times X$ matrix A is the derivative of the first XS entries with respect to $\mathbf{u}$

$$
A = \begin{bmatrix}
h''(u_1)\mathbf{P}_1^3 & 0 & \dots & 0 \\
0 & h''(u_2)\mathbf{P}_2^3 & \dots & 0 \\
\vdots & \vdots & & \vdots \\
0 & 0 & \dots & h''(u_X)\mathbf{P}_X^3
\end{bmatrix},
$$

where $\mathbf{P}_x^e = \left(p_{x,1}^e,...,p_{x,S}^e\right)'$. The $3 \times X$ matrix B includes the derivatives of the last 3 equations (IR and ICs) with respect to $\mathbf{u}$:

$$
B = \begin{bmatrix}
K(3)\mathbf{P}_1^3 \cdot \mathbf{1}_S & K(3)\mathbf{P}_2^3 \cdot \mathbf{1}_S & \dots & K(3)\mathbf{P}_X^3 \cdot \mathbf{1}_S \\
K(2)\mathbf{P}_1^2 \cdot \mathbf{1}_S & K(2)\mathbf{P}_2^2 \cdot \mathbf{1}_S & \dots & K(2)\mathbf{P}_S^2 \cdot \mathbf{1}_S \\
K(1)\mathbf{P}_1^1 \cdot \mathbf{1}_S & K(1)\mathbf{P}_2^1 \cdot \mathbf{1}_S & \dots & K(1)\mathbf{P}_S^1 \cdot \mathbf{1}_S
\end{bmatrix}, \tag{16}
$$

14

where $\mathbf{1_S} \equiv (1, 1, ..., 1)$ is the vector of ones with length S. The $XS \times 3$ matrix C is the derivative of the the first XS equations with respect to the multipliers:

$$
C = \begin{bmatrix}
K(1)p_{1,1}^1 & K(2)p_{1,1}^2 & -K(3)p_{1,1}^3 \\
K(1)p_{1,2}^1 & K(2)p_{1,2}^2 & -K(3)p_{1,2}^3 \\
\vdots & & \\
K(1)p_{1,S}^1 & K(2)p_{1,S}^2 & -K(3)p_{1,S}^3 \\
\vdots & & \\
K(1)p_{X,1}^1 & K(2)p_{X,1}^2 & -K(3)p_{X,1}^3 \\
K(1)p_{X,2}^1 & K(2)p_{X,2}^2 & -K(3)p_{X,2}^3 \\
\vdots & & \\
K(1)p_{X,S}^1 & K(2)p_{X,S}^2 & -K(3)p_{X,S}^3
\end{bmatrix}. \tag{17}
$$

and $\mathbf{0}_{3\times3}$ is a 3×3 null matrix corresponding to the derivative of the last 3 equations (IR and ICs) with respect to the multipliers.

The derivative of the first XS equations with respect to $\{G(3), G(2), G(1)\}$ is the $XS \times 3$ null matrix $\mathbf{0}_{XS\times3}$. The derivative of the last 3 equations (IR and ICs) with respect to $\{G(3), G(2), G(1)\}$ is the 3×3 identity matrix $\mathbf{I}_3$. Thus, if $\mathbf{K}$ is constant, $\Theta = \mathbf{G}$, and we have $D_\Theta = D_\mathbf{G} = \mathbf{0}_{XS\times3}$, and $E_\Theta = E_\mathbf{G} = \mathbf{I}_3$.

The derivative of the first XS equations with respect to $\{K(3), K(2), K(1)\}$ is

$$
D_\mathbf{K} = \begin{bmatrix}
-\lambda p_{1,1}^3 & \mu_2 p_{1,1}^2 & \mu_1 p_{1,1}^1 \\
-\lambda p_{1,2}^3 & \mu_2 p_{1,2}^2 & \mu_1 p_{1,2}^1 \\
\vdots & & \\
-\lambda p_{1,S}^3 & \mu_2 p_{1,S}^2 & \mu_1 p_{1,S}^1 \\
\vdots & & \\
-\lambda p_{X,1}^3 & \mu_2 p_{X,1}^2 & \mu_1 p_{X,1}^1 \\
-\lambda p_{X,2}^3 & \mu_2 p_{X,2}^2 & \mu_1 p_{X,2}^1 \\
\vdots & & \\
-\lambda p_{X,S}^3 & \mu_2 p_{X,S}^2 & \mu_1 p_{X,S}^1
\end{bmatrix}.
$$

The derivative of the last 3 equations with respect to $\{K(3), K(2), K(1)\}$ is

$$
E_\mathbf{K} = \begin{bmatrix}
\sum_{x=1}^X u_x \sum_s p_{x,s}^3 & 0 & 0 \\
0 & \sum_{x=1}^X u_x \sum_s p_{x,s}^2 & 0 \\
0 & 0 & \sum_{x=1}^X u_x \sum_s p_{x,s}^1
\end{bmatrix}.
$$

15

Thus, if $\mathbf{G}$ is constant, $\Theta = \mathbf{K}$, and we have $D_\Theta = D_\mathbf{K}$, and $E_\Theta = E_\mathbf{K}$.

Next, we calculate the derivative with respect to the probabilities. The derivative of the first XS equations with respect to $(p^3_{x,s})$ is the $XS \times XS$ matrix:

$$H_3 = \begin{bmatrix} [h'(u_1) - K(3)\lambda]\,\mathbf{I}_S & \mathbf{0}_{S\times S} & \dots & \mathbf{0}_{S\times S} \\ \mathbf{0}_{S\times S} & [h'(u_2) - K(3)\lambda]\,\mathbf{I}_S & \dots & \mathbf{0}_{S\times S} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{S\times S} & \mathbf{0}_{S\times S} & \dots & [h'(u_X) - K(3)\lambda]\,\mathbf{I}_S \end{bmatrix},$$

where $\mathbf{I}_S$ is the $S \times S$ identity matrix. The derivative of the last three equations with respect to $(p^3_{x,s})$ is:

$$J_3 = \begin{bmatrix} u_1 K(3)\mathbf{1}_S & \dots & u_X K(3)\mathbf{1}_S \\ \mathbf{0}_S & \dots & \mathbf{0}_S \\ \mathbf{0}_S & \dots & \mathbf{0}_S \end{bmatrix},$$

which is a $3 \times XS$ matrix and, as before, $\mathbf{1}_S$ is the row vectors of ones with length S.

Proceeding in a similar way with respect to $(p^2_{x,s})$ and $(p^1_{x,s})$, gives

$$H_2 = \begin{bmatrix} \mu_2 K(2)\mathbf{I}_S & \mathbf{0}_{S\times S} & \dots & \mathbf{0}_{S\times S} \\ \mathbf{0}_{S\times S} & \mu_2 K(2)\mathbf{I}_S & \dots & \mathbf{0}_{S\times S} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{S\times S} & \mathbf{0}_{S\times S} & \dots & \mu_2 K(2)\mathbf{I}_S \end{bmatrix} = \mu_2 \mathbf{I}_{SX},$$

$$J_2 = \begin{bmatrix} \mathbf{0}_S & \dots & \mathbf{0}_S \\ u_1 K(2)\mathbf{1}_S & \dots & u_X K(2)\mathbf{1}_S \\ \mathbf{0}_S & \dots & \mathbf{0}_S \end{bmatrix},$$

and

$$H_1 = \mu_1 K(1)\mathbf{I}_{SX},$$

$$J_1 = \begin{bmatrix} \mathbf{0}_S & \dots & \mathbf{0}_S \\ \mathbf{0}_S & \dots & \mathbf{0}_S \\ u_1 K(1)\mathbf{1}_S & \dots & u_X K(1)\mathbf{1}_S \end{bmatrix}.$$

Note that $DF_\mathbf{P} = \begin{bmatrix} H_3 & H_2 & H_1 \\ J_3 & J_2 & J_1 \end{bmatrix}$ has $XS+3$ rows and $3XS$ columns. Since $XS+3 < 3XS$, it suffices to show that $DF_\mathbf{P}$ has full row rank: for any $\mathbf{y} \in \mathbb{R}^{XS+3}$,

$$\underbrace{\mathbf{y}}_{1\times(XS+3)} \times \underbrace{DF_\mathbf{P}}_{(XS+3)\times 3XS} = \underbrace{\mathbf{0}}_{1\times 3XS} \implies \mathbf{y} = \underbrace{\mathbf{0}}_{1\times(XS+3)}.$$

16

Let $DF_{\mathbf{P}_i} = \begin{bmatrix} H_i \\ J_i \end{bmatrix}$. First, expanding $\mathbf{y} \times DF_{\mathbf{P}_2} = \mathbf{0}$ gives:

$$x_1\mu_2 K(2) + x_{SX+2}u_1 K(2) = 0,$$
$$\vdots$$
$$x_S\mu_2 K(2) + x_{SX+2}u_1 K(2) = 0,$$
$$x_{S+1}\mu_2 K(2) + x_{SX+2}u_2 K(2) = 0,$$
$$\vdots$$
$$x_{2S}\mu_2 K(2) + x_{SX+2}u_2 K(2) = 0,$$
$$\vdots$$
$$x_{S(X-1)+1}\mu_2 K(2) + x_{SX+2}u_X K(2) = 0,$$
$$\vdots$$
$$x_{SX}\mu_2 K(2) + x_{SX+2}u_X K(2) = 0,$$

which implies

$$\mu_2 K(2)y_1 + u_1 K(2)y_{XS+2} = \mu_2 K(2)y_2 + u_1 K(2)y_{XS+2}$$
$$= ... = \mu_2 K(2)y_S + u_1 K(2)y_{XS+2} = 0$$
$$\mu_2 K(2)y_{S+1} + u_2 K(2)y_{XS+2} = \mu_2 K(2)y_{S+2} + u_2 K(2)y_{XS+2}$$
$$= ... = \mu_2 K(2)y_{2S} + u_2 K(2)y_{XS+2} = 0$$
$$\vdots$$
$$\mu_2 K(2)y_{S(X-1)+1} + u_X K(2)y_{XS+2} = \mu_2 K(2)y_{S(X-1)+2} + u_X K(2)y_{XS+2}$$
$$= ... = \mu_2 K(2)y_{SX} + u_X K(2)y_{XS+2} = 0.$$

Dividing through by $K(2) > 0$ and rearranging gives:

$$\mu_2 y_1 = \mu_2 y_2 = ... = \mu_2 y_S = -u_1 y_{XS+2} \tag{18}$$
$$\mu_2 y_{S+1} = \mu_2 y_{S+2} = ... = \mu_2 y_{2S} = -u_2 y_{XS+2}$$
$$\vdots$$
$$\mu_2 y_{S(X-1)+1} = \mu_2 y_{S(X-1)+2} = ... = \mu_2 y_{SX} = -u_X y_{XS+2}.$$

Similarly, expanding $\mathbf{y} \times DF_{\mathbf{P}_1} = \mathbf{0}$, yields

$$
\begin{aligned}
\mu_1 K(1)y_1 &= \mu_1 K(1)y_2 = ... = \mu_1 K(1)y_S = -u_1 K(1)y_{XS+3} \quad (19) \\
\mu_1 K(1)y_{S+1} &= \mu_1 K(1)y_{S+2} = ... = \mu_1 K(1)y_{2S} = -u_2 K(1)y_{XS+3} \\
&\quad\vdots \\
\mu_1 K(1)y_{S(X-1)+1} &= \mu_1 K(1)y_{S(X-1)+2} = ... = \mu_1 K(1)y_{SX} = -u_X K(1)y_{XS+3}.
\end{aligned}
$$

with $K(1) > 0$. Recall that $\mu_1 \geq 0$ and $\mu_2 \geq 0$ and at least one of them is strict. Thus,

$$
\begin{aligned}
y_1 &= y_2 = ... = y_S = \bar{y}^1 \\
y_{S+1} &= y_{S+2} = ... = y_{2S} = \bar{y}^2 \\
&\quad\vdots \\
y_{S(X-1)+1} &= y_{S(X-1)+2} = ... = y_{XS} = \bar{y}^X.
\end{aligned}
$$

From equation (18), we have:

$$
\begin{aligned}
\mu_2 \bar{y}^1 &= -u_1 y_{XS+2} \\
&\quad\vdots \qquad\qquad\qquad\qquad\qquad (20) \\
\mu_2 \bar{y}^X &= -u_X y_{XS+2}.
\end{aligned}
$$

Second, recall that $DF_{(\mu_1,\mu_2,\lambda)} = \begin{bmatrix} C \\ \mathbf{0}_{3\times 3} \end{bmatrix}$, where C is described in (17). Thus, $y \times DF_{(\mu_1,\mu_2,\lambda)} = \mathbf{0}$ gives

$$
\sum_{x,s} \bar{y}^x K(1)p^1_{x,s} = 0, \quad \sum_{x,s} \bar{y}^x K(2)p^2_{x,s} = 0, \quad \sum_{x,s} \bar{y}^x K(3)p^3_{x,s} = 0 \quad \forall x. \quad (21)
$$

Multiplying both sides of the first equation in (21) by $\mu_2 \geq 0$:

$$
\mu_2 \sum_{x,s} \bar{y}^x K(1)p^1_{x,s} = K(1) \sum_{x,s} (\mu_2 \bar{y}^x) p^1_{x,s} = 0. \quad (22)
$$

However, from equation (20), we have

$$
K(1) \sum_{x,s} (\mu_2 \bar{y}^x) p^1_{x,s} = -y_{XS+2} K(1) \sum_{x,s} u_x p^1_{x,s} = -y_{XS+2}(\bar{U} - G(1)), \quad (23)
$$

where the last equality follows from the IC associated with $e = 1$. Let $G(1) \neq \bar{U}$ (the set of parameters for which $\bar{U} = G(1)$ have zero Lebesgue measure). Then, (22) and

(23) imply $y_{XS+2} = 0$. Applying the same logic to the second equation in (21) gives $y_{XS+3} = 0$.

Third, recall from equations (18) and (19) that, for all x,

$$\mu_2 \bar{y}^x = -u_x y_{XS+2} \quad \text{and} \quad \mu_1 \bar{y}^x = -u_x y_{XS+3}.$$

Moreover, $\mu_1 \geq 0$ and $\mu_2 \geq 0$ with one of them strict. Given $y_{XS+2} = y_{XS+3} = 0$, it then follows that $\mu_1 \bar{y}^x = \mu_2 \bar{y}^x = 0$, which, because either $\mu_1 \neq 0$ or $\mu_2 \neq 0$, implies $\bar{y}^x = 0$ for all x.

Fourth, expanding $\mathbf{y} \times DF_{\mathbf{P}_3} = \mathbf{0}$ gives:

$$y_1 \left[h'(u_1) - K(3)\lambda \right] + y_{SX+1} u_1 K(3) = 0,$$

$$\vdots$$

$$y_S \left[h'(u_1) - K(3)\lambda \right] + y_{SX+1} u_1 K(3) = 0,$$

$$y_{S+1} \left[h'(u_2) - K(3)\lambda \right] + y_{SX+1} u_2 K(3) = 0,$$

$$\vdots$$

$$y_{2S} \left[h'(u_2) - K(3)\lambda \right] + y_{SX+1} u_2 K(3) = 0,$$

$$\vdots$$

$$y_{S(X-1)+1} \left[h'(u_X) - K(3)\lambda \right] + y_{SX+1} u_X K(3) = 0,$$

$$\vdots$$

$$y_{SX} \left[h'(u_X) - K(3)\lambda \right] + y_{SX+1} u_X K(3) = 0.$$

Given that $y_1 = y_2 = \cdots = y_{SX} = 0$ and $K(3) > 0$, this implies that either $u_1 = u_2 = \cdots = u_X$ or $y_{SX+1} = 0$. However, the former is impossible: either such a contract violates at least one IC, or it satisfies all ICs. In the latter case, a constant wage would induce e^*, which was ruled out. It follows that $y_{SX+1} = 0$. Therefore, $\mathbf{y} \times DF_{\mathbf{P}} = \mathbf{0} \implies \mathbf{y} = \mathbf{0}$, showing that $DF_{\mathbf{P}}$ has full row rank.