Relational Incentive Contracts and Performance Measurement

Chang Koo Chi[1] and Trond E. Olsen*[2]

[1]Department of Economics, Norwegian School of Economics
[2]Department of Business and Management Science, Norwegian School of Economics

January 16, 2019

Abstract

This paper analyzes relational contracts under moral hazard. We first show that if the available information (signal) about effort satisfies a novel condition, the monotone likelihood ratio *transformation* property, then irrespective of whether the first-order approach (FOA) is valid, the optimal bonus scheme takes a simple form. The scheme rewards the agent a fixed bonus if the signal's likelihood ratio exceeds a threshold, but in contrast to the FOA contract characterized in Levin (2003), the threshold is not necessarily zero. We next derive a sufficient and necessary condition for non-verifiable information to improve the efficiency of a contract. Our new informativeness criterion sheds light on the nature of an ideal performance measure in relational contracting.

JEL CLASSIFICATION: D82, D86

KEYWORDS: Relational contracts; non-verifiable performance measures; first-order approach; bonus scheme; informativeness criterion

1. Introduction

In many organizations, managerial incentives are frequently implicit. Recent empirical studies report that, since the 1990s firms have increasingly been adopting a practice of using non-financial measures such as customer satisfaction scores, leadership, or other subjective evaluations, to assess and pay for managerial performance.[1] The use of non-verifiable measures in a contract sup-

*We would like to thank Malin Arve, Jonathan Bonham, Tilman Börgers, In-Koo Cho, Jinwoo Kim, David Miller, Joel Watson, and participants in seminars at the Hanyang University, the University of Michigan, the University of Oslo, and the University of San Diego, as well as participants at the 2018 Asian Econometric Society Meetings and the 13th EIASM Workshop on Accounting and Economics for helpful comments and encouragement.
Email addresses: chang-koo.chi@nhh.no (Chi); trond.olsen@nhh.no (Olsen)

[1]For instance, Murphy and Oyer (2003) and Gillan, Hartzell and Parrino (2009) found that more than one-half of their sample firms base employees' annual bonus at least in part on non-financial measures of individual performance.

plements the weakness of objective measures, but the contract cannot be enforced by external parties. Nevertheless, if contracting parties repeatedly transact over time, a wide array of contracts can be self-enforced by the value of an ongoing relationship. Such relational contracts between firms were observed several decades ago by legal scholars (Macaulay (1963)), and have since been extensively analyzed and applied in economics and other areas.[2] However, the related literature has mainly focused on the problem of designing an optimal contract with non-verifiable information (i.e., how to *pay*) and paid little attention to the problem of choosing an ideal performance measure among many alternatives (how to *evaluate*), although both appropriate measures and well-designed contracts are key ingredients to successful long-term relations.

This paper addresses both of these two issues in relational contracting. We consider a standard repeated agency model in which two risk-neutral parties interact over time. In each period, an agent exerts hidden effort that creates a surplus to be shared, and a principal incentivizes the agent through a contract. We assume that all available performance measures (or signals) are imperfect and non-verifiable, but commonly observable to both parties.[3] We formulate this agency problem as a two-stage mechanism: the principal first chooses a performance measurement system and then offers an incentive contract based on the chosen system. By virtue of Levin (2003), our analysis of optimal contracts focuses on the stationary contract that consists of a time-invariant base salary and discretionary bonus pay. Within this environment, we characterize an optimal stationary contract and provide a novel criterion for one evaluation system to be more informative than another in the spirit of Holmström (1979).

The main contribution of this paper is thus two-fold. We first present a new approach to solving relational contract problems under moral hazard. Our approach complements the standard analysis that has been limited to the first-order approach (FOA), and can be applied to a wide class of multi-signal problems where the local approach cannot be justified. To be specific, we provide a condition on the measurement system that we term the monotone likelihood ratio *transformation* property (MLRTP). We then show that as long as the available system satisfies this condition, the optimal bonus scheme retains a simple hurdle structure as in the FOA contract characterized by Levin (2003), whether the FOA is valid or not.[4] Our characterization is of practical interest in its own right, but also allows us to relax several of the assumptions on the primitives that have been made to validate the FOA.

We next make use of this result to investigate the principal's problem of choosing a performance measurement system: Between two (multivariate) measurement systems, which one en-

[2]Seminal contributions include Klein and Leffler (1981), Bull (1987), MacLeod and Malcomson (1989) and Levin (2003). See also Malcomson (2012) for a review.

[3]Much non-verifiable information in practice is not observed by the agent, in particular when information is gathered through the principal's subjective appraisal. A standard example of non-verifiable but observable measures in organizations is a performance evaluation by other human resource divisions or customer's satisfaction scores. Hence we abstract away other problems of subjective measures such as leniency bias (MacLeod (2003)), favoritism (Prendergast and Topel (1996)), or influence activities (Milgrom (1988)).

[4]To clarify our first contribution, our aim is not to provide a condition that ensures validity of the FOA but to provide a condition under which the optimal contract takes a simple hurdle form like the FOA contract. A recent paper by Hwang (2016) develops a condition in the same environment as ours under which the FOA is justified.

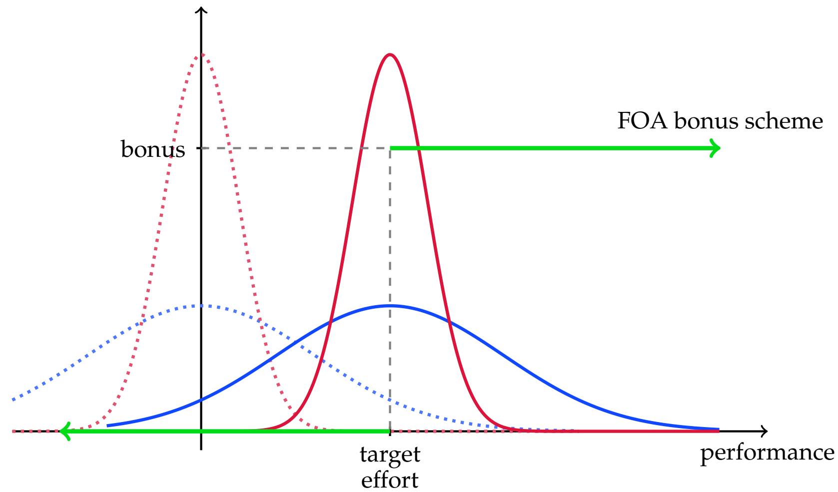


Figure 1: First order approach failure

ables the parties to attain a higher surplus through the relational contract and thus realize a more successful relationship? Invoking the simple structure of the optimal contract and its applicability to a wide range of contracting environments, we establish a novel informativeness criterion: the *likelihood ratio order*.

To illustrate how these two results are interrelated, suppose that the agent's performance $X = e + \epsilon$ is a univariate signal and is subject to his effort e and the additive noise ϵ with a normal distribution $N(0, \sigma^2)$.[5] Under this specification, it is natural to presume that as the standard deviation (σ) of the noise gets lower, the principal could effectively reduce the agency cost and hence elicit higher effort from the agent. In fact, under the assumption that the FOA is valid, a simple analysis confirms this presumption. However, this local approach does not work when σ is sufficiently small, in that the optimal effort derived from the FOA is a stationary point of the agent's expected utility, but does not maximize his utility.[6] As a result, the agent does not exert the target effort but deviates by choosing a low effort that is distant from the target. A characterization of the optimal contract has been lacking for this case, and a ranking of measurement systems based on the FOA has therefore been incomplete even in this most natural setting. That is, even if one system is a garbling of another in the sense of Blackwell (1951, 1953), the existing approach cannot tell which one is more informative in relational contracting.

[5]Payments are bounded due to self-enforcement constraints, and the type of schemes with arbitrarily large penalties suggested by Mirrlees (1979) to approach the first best are not feasible here.

[6]Kvaløy and Olsen (2014) pointed out that FOA is valid only if the output shock is sufficiently diffuse in this setting. The reason for this can be easily seen in Figure 1. Given the FOA bonus scheme (the green-colored line), when σ is relatively small as depicted by the red curve, the agent's marginal gain from effort at the target effort is very high. On the other hand, the marginal gain at low effort distant from the target is very low. That is, the FOA contract provides the agent with high-powered (marginal) incentives but does not necessarily maximize the agent's expected payoff at the target effort. A similar discussion also can be found in the tournament literature stemming from Lazear and Rosen (1981). For the same reason, unless the shock to individual output is sufficiently diffuse, the objective function of each agent is not globally concave, and so FOA may fail.

We fill this gap by developing an alternative approach to characterizing the optimal contract. Our approach does not call for the so-called Mirrlees-Rogerson conditions on the primitives and thus can be applied to a large family of signals, even multivariate ones.[7] In Section 3 we introduce a sufficient condition, the MLRTP, for our approach to be valid. The condition endows the likelihood ratios of the signal with an ordinal property, allowing the principal to evaluate the agent's performance on the basis of the ratios for payment of a bonus. The MLRTP is a weaker condition than the monotone likelihood ratio property in the one-signal case. Hence our approach can be applied to the example above, even when σ is small. The first main result shows that as long as the measurement system satisfies the MLRTP, the optimal bonus scheme takes a hurdle form for the likelihood ratio as is displayed in Figure 1: the agent is awarded a bonus if a relevant performance measure - the likelihood ratio – clears a hurdle.

In contrast with the FOA contract, however, the optimal hurdle is no longer necessarily set at zero. We demonstrate that when the FOA is invalid, the optimal hurdle reflects a trade-off between providing on the one hand strong incentives for effort on the margin (locally) and curbing on the other hand global deviations to distinctly lower effort. Depending on the agent's inclination to deviate from the target effort, the hurdle has to be adjusted accordingly. In Section 3, we revisit the example above and illustrate how this trade-off determines the optimal hurdle. It turns out that when σ is sufficiently small, the optimal contract features a negative hurdle, put another way, a more lenient one than the FOA contract.[8] Furthermore, the optimal contract equipped with an adjusted hurdle does implement higher effort as σ decreases. Therefore, our approach provides not only a full characterization of optimal contracts, but a complete (and intuitive) ranking of available measurement systems in the example.

The example above also suggests that, like verifiable measures used in explicit contracts, non-verifiable ones in relational contracts can be ranked by a standard statistical order. It is intuitive that an improvement of the measurement system in the sense of Blackwell garbling alleviates the agency problem and results in a more efficient contract. However, there is a noteworthy difference between these two types of contracts. While the agency costs arise from the constrained monetary incentives due to the agent's risk aversion in explicit contracts, the costs arise from the constrained incentives due to the enforcement problem in relational contracts. Given these different sources of agency costs, it is unclear whether the existing informativeness criterions in agency models, for example the sufficient statistic theorem in Holmström (1979), are suited for ranking non-verifiable signals.

In Section 4 we introduce a new criterion, the likelihood ratio order, which provides a tight condition for one measurement system to be more informative than another in relational contracts. Our criterion rests on the distribution of the signal's likelihood ratio, which follows naturally from the fact that the ratio plays a role as a main performance index in the optimal contract. As

[7] In the formal contracting with multivariate verifiable measures, Conlon (2009) and Jung and Kim (2015) identify a set of conditions under which the FOA is justified. See Kirkegaard (2017a) for another approach using stochastic orders.

[8] We provide sufficient conditions under which the optimal bonus scheme exhibits a nonpositive (more lenient) or nonnegative (more strict compared to the FOA contract) hurdle in a general framework.

the ratio is a unidimensional information variable, the criterion provides a unified treatment for a comparison of multivariate (noninclusive) signals satisfying the MLRTP. Simply put, the stochastic order determines a partial ranking between two signals based on the variability of likelihood ratios. If one signal's likelihood ratio is more variable with regard to the agent's choice of effort than another, then the signal conveys more information about his potential deviations. Hence the principal can more effectively control the hidden actions by designing a bonus plan based on that signal. It turns out that the criterion is also necessary for the principal to induce higher effort from the agent.[9] Consequently, our second result provides a full characterization for the notion of informativeness in relational contracts.

We show that our informativeness criterion is closely related to the notion of precision introduced by Lehmann (1988). To be specific, Lehmann's criterion applied to the likelihood ratios of the signals under consideration provides a sufficient condition for the likelihood ratio order. Compared to Blackwell's garbling (or sufficiency), Lehmann's criterion is not just easier to check, but also provides a link to the existing information rankings developed in agency theory. The link sheds light on how ideal performance measures differ between explicit and relational contracts.

Related Literature

This paper is related to two strands of literature in contract theory, in that it develops an alternative approach for the optimal design of incentive contracts, and provides a new criterion for an ideal performance measure in relational contracting environments.

Our first main result on optimal bonus schemes in relational contracting complements the seminal work by Levin (2003), which characterizes an optimal contract in the environment where the FOA is valid and the univariate performance measure is exogenously given by the principal's objective (output). A recent paper by Hwang (2016) allows the principal to use alternative multivariate measures and proposes a sufficient condition under which the agent's problem is globally concave with respect to his effort and thus the FOA is justified. Our approach is different from his in the aspect that instead of conditions that justify the FOA, we seek for conditions that ensure the optimal bonus scheme to take a simple form.[10] In the same spirit as this paper, Poblete and Spulber (2012) analyze a static model of financial contracting between two risk-neutral parties but with two-sided limited liability, and provide a condition under which debt-style contracts are optimal regardless of validity of the FOA. As has been pointed out by Levin (2003), the enforcement problem in relational contracts imposes a lower and upper bound for monetary incentives, much like limited liability does. In Appendix B we further discuss and compare the analysis in Poblete

[9] More precisely, the necessary part can be established by showing that if one signal (say X) does not dominate another (Y) in the likelihood ratio order, there exist a pair of contracting parties, represented by the principal's objective and the agent's cost function from effort, for whom the principal prefers to design an incentive contract based on Y rather than X.

[10] It is worthwhile to note that MLRTP is complementary to the condition of Hwang (2016), the local convexity of distribution function condition (LCDFC). As we will discuss in Section 2, when the additive noise ϵ has a small variance, the signal $X = e + \epsilon$ does not satisfy his condition but obeys ours. On the other hand, there is a set of signals satisfying LCDFC but not MLRTP.

and Spulber (2012) with ours.

Our second result on performance measurement extends a line of research initiated by Holmström (1979). The related literature is mostly restricted to verifiable signals in the standard formal contracting problem with a risk-averse agent.[11] The classic results, including Holmström (1979), Gjesdal (1982) and Grossman and Hart (1983), were developed by applying Blackwell's theorem. Kim (1995) subsequently showed that provided the FOA is valid, the signal having a more dispersed likelihood ratio distribution (in terms of mean-preserving spread) is more informative in explicit contracts.[12] Our informativeness criterion has a similar flavor to Kim's in that both criteria pertain to the variability of likelihood ratios and thus provide a unified treatment of comparison of signals regardless of their dimension. In addition to different notions of variability, one notable difference is that Kim's criterion is based on the variability of the ratio in response to the agent's local deviations, whereas our criterion is on the variability in response to all possible downward deviations. This highlights the different sources of agency costs in formal and relational contracts.

The rest of this paper is organized as follows. In Section 2 we present the model and formulate the optimal stationary contract problem. In Section 3 we introduce the condition MLRTP, illustrate its implications, and characterize the optimal bonus scheme. In Section 4 we examine the problem of choosing an ideal performance measure. Section 5 concludes. All omitted proofs are relegated to Appendix A and more details on the MLRTP can be found in Appendix B.

2. The Model

We consider a repeated moral hazard model between a risk-neutral principal and agent, as in Levin (2003). The two parties interact in discrete time over an infinite time horizon with a common discount factor $\delta \in (0,1)$. At the outset of each period $t = 1, 2, \cdots$, the principal offers the agent a compensation package that consists of a base salary and a bonus scheme. The agent, if he accepts the offer, privately chooses a level of effort e_t from $[0, \bar{e}] \subset \Re$ by incurring a cost of $c(e_t)$. If he rejects, then nothing happens until the next period, and the principal and the agent obtain a reservation payoff of $\overline{\pi}$ and $\overline{u}$, respectively.

The agent's effort e_t produces an expected profit of $v(e_t)$ for the principal, and generates a set of commonly observable but unverifiable outcomes $\mathbf{x}_t = (x_t^1, \cdots, x_t^n)$.[13] The cumulative distribution function (CDF) of outcomes is $F(\cdot, e_t)$ with support $\mathbb{X} \subset \Re^n$. We call the outcome-generating process $\mathbf{X}_t \sim F(\cdot, e_t)$ a *signal* hereafter.[14] We use a capital letter for a random variable and a

[11]To our best knowledge, one exception is the paper by Dewatripont, Jewitt and Tirole (1999) which compares the market signals about the agent's unknown talent in the career concern model. Their paper finds that an improvement of signals (even in the sense of Blackwell sufficiency) may strengthen or undermine incentives to work.

[12]A recent paper by Chi and Choi (2018) establishes that Kim's mean-preserving spread criterion is also necessary for a verifiable measure to be more informative in formal contracts, under the assumption that the FOA is valid. Their paper also shows that for univariate signals satisfying the monotone likelihood ratio property, Kim's criterion is equivalent to the Lehmann (1988) order.

[13]In standard agency models with a univariate signal, the principal's objective is exogenously given by the expected value of the signal; i.e. $v(e_t) = E(X_t|e_t)$. In our model, the realized benefit in period t need not be part of $\mathbf{x}_t$, that is, the exact benefit may or may not be observed by both parties when the bonus is paid. We discuss more details in Section 4.

[14]A signal is therefore defined by a set of cumulative distribution functions $\{F(\mathbf{x}, e),\ e \in [0, \bar{e}]\}$. In contract theory

small letter for its realization. Also, a bold letter indicates a vector and a normal letter indicates a scalar. We denote by $\omega \equiv \langle v, c, \overline{\pi}, \overline{u}, \delta \rangle$ the 5-tuple elements that describe a pair of parties' characteristics. In conjunction with the signal $\mathbf{X}$, ω defines an agency problem of our interest. Throughout the paper, we make the following assumptions on the environment:

ASSUMPTION 1.

(A1) *The principal's expected profit v and the agent's cost function from effort are increasing and continuously differentiable over $[0, \bar{e}]$.*

(A2) *The net per-period surplus from forming a relationship,*

$$s(e) \;=\; v(e) - c(e) - \overline{\pi} - \overline{u},$$

is increasing on $[0, e^{FB}]$ where $e^{FB} \equiv \operatorname*{argmax}_{e \in [0,\bar{e}]} s(e)$ denotes the surplus-maximizing (first-best) *effort.*

(A3) $s(0) < 0 < s(e^{FB})$.

(B) *The CDF $F(\mathbf{x}_t, e_t)$ is twice continuously differentiable with respect to both arguments, and we denote by $f(\mathbf{x}_t, e_t)$ the density function.*

After observing a realization $\mathbf{x}_t$ of the signal, the principal pays the fixed salary w_t as promised and decides which bonus β_t to pay. Here w_t is a legally enforceable payment that the principal can commit to, whereas $\beta_t : \mathbb{X} \to \Re$ is a discretionary payment. That is, the bonus represents the parties' coordinated behavior to reward or punish each other depending on the observed performance. The principal now obtains a payoff of the realized profit less the total payment $w_t + \beta_t(\mathbf{x}_t)$, and the agent obtains a payoff of $w_t + \beta_t(\mathbf{x}_t) - c(e_t)$. Each party then decides whether to continue their relationship or separate. If at least one party decides to walk away, the entire game ends and each party obtains the reservation payoff from the next period on.

An optimal contract in this environment can be taken to be *stationary* as is shown by Levin (2003), which greatly simplifies the problem. In a stationary contract, the principal proposes the same base salary $w_t = w$ and bonus scheme $\beta_t = \beta$ every period, in anticipation that such payments induce the agent to choose effort $e_t = e$. The main intuition behind this result comes from the fact that the two instruments for incentives, the promised utility to the agent and the bonus scheme, are equally effective under risk-neutrality (and absence of limited liability constraints). Accordingly, we can focus on such a stark form of contract in which the agent's continuation utility remains constant over time and incentives are provided by the instantaneous bonus only. Dropping the time index, we represent a stationary contract by (w, β, e).

For a contract (w, β, e) to be sustainable, its implicit part (β, e) must respect the following two conditions. First, the bonus scheme should provide the agent with a proper incentive, so that the desired effort e must maximize the agent's expected payoff. Since the base salary is immaterial

literature, the signal is often referred to as a performance measurement system or an information system.

to the agent's incentive, we abstract away w and write the agent's expected payoff from choosing effort e' as $u(\beta, e') \equiv \mathbb{E}[\beta(\mathbf{X})|e'] - c(e')$. With this expression, the global incentive compatibility (IC) condition can be expressed as

$$u(\beta, e) \;\geq\; u(\beta, e') \quad \forall\, e' \in [0, \bar{e}]. \tag{IC$_\text{G}$}$$

In addition, the bonus scheme must be self-enforcing because the parties have no legal obligation to pay β. A bonus will be paid as promised only if both parties wish so, put another way, only if the expected payoffs from holding onto the payment rule to each party are higher than those from reneging on the rule. Assuming that each party responds by terminating future transactions to breach of contracts, we can formulate the self-enforcement condition as the following inequality: for all $\mathbf{x} \in \mathbb{X}$,

$$\frac{\delta}{1-\delta}\Big(\overline{u} - w - u(\beta, e)\Big) \;\leq\; \beta(\mathbf{x}) \;\leq\; \frac{\delta}{1-\delta}\Big(\pi(w, \beta, e) - \overline{\pi}\Big),$$

where $\pi(w, \beta, e) \equiv v(e) - w - \mathbb{E}[\beta(\mathbf{X})|e]$ denotes the principal's expected net (per-period) profit from the stationary contract. Since the base salary allows us to separate the problem of efficient contracting from the problem of distribution (Theorem 1 in Levin (2003)), we may assume without loss of generality that the agent's participation constraint is binding. Hence we take $w = \overline{u} - u(\beta, e)$, which further simplifies the enforcement condition into

$$0 \;\leq\; \beta(\mathbf{x}) \;\leq\; \frac{\delta}{1-\delta} s(e) \quad \forall\, \mathbf{x} \in \mathbb{X}. \tag{SE}$$

An optimal contract (w^*, β^*, e^*) is then characterized by a solution to the constrained optimization problem: max $s(e)$ subject to (IC$_\text{G}$) and (SE). The standard approach to this problem is to replace the global constraint (IC$_\text{G}$) with a local stationary condition that prevents the agent from deviating to other effort nearby e^*, and then check if the obtained solution is indeed optimal. This procedure, the so-called first-order approach (FOA), implicitly assumes that the following local condition is sufficient for (IC$_\text{G}$):

$$\frac{\partial u(\beta, e^*)}{\partial e} \;=\; \int \beta(\mathbf{x}) l(\mathbf{x}, e^*) f(\mathbf{x}, e^*) d\mathbf{x} \;-\; c'(e^*) \;=\; 0, \tag{IC$_\text{L}$}$$

where $l(\mathbf{x}, e^*) \equiv f_e(\mathbf{x}, e^*)/f(\mathbf{x}, e^*)$ is the likelihood ratio at e^* (the subscript of f denotes the partial derivative). The information variable $l(\mathbf{x}, e^*)$ captures how likely it is that the agent has made the desired effort e^* rather than other nearby effort given $\mathbf{x}$.

Replacing (IC$_\text{G}$) with (IC$_\text{L}$), the associated Lagrangian becomes linear in β. As a result, we obtain the FOA bonus scheme $\beta^\dagger$ that exhibits the bang-bang property: $\beta^\dagger(\mathbf{x}) = 0$ if $l(\mathbf{x}, e^\dagger) < 0$ and $\beta^\dagger(\mathbf{x}) = b^\dagger \equiv \frac{\delta}{1-\delta} s(e^\dagger)$ if $l(\mathbf{x}, e^\dagger) \geq 0$. We use the different superscript "$\dagger$" for the FOA contract, in order to distinguish from the optimal contract. This standard approach is based on the local constraint (IC$_\text{L}$) which precludes the agent's possible local deviations from $e^\dagger$ only. Hence the FOA contract is designed so as to provide the strongest marginal incentives at the target effort, given the self-enforcement constraint (SE).

To conclude that $(w^\dagger, \beta^\dagger, e^\dagger)$ constitutes an optimal contract, we need to verify that $(\beta^\dagger, e^\dagger)$ satisfies the global IC constraint. Using the form of $\beta^\dagger$, (IC_G) can be rewritten as

$$b^\dagger \Pr(l(\mathbf{x}, e^\dagger) > 0 \,|\, e) - c(e) \;\leq\; b^\dagger \Pr(l(\mathbf{x}, e^\dagger) > 0 \,|\, e^\dagger) - c(e^\dagger), \quad \forall\, e \in [0, \bar{e}]. \tag{1}$$

Observe that if the expression on the left-hand side of (1) is globally concave in e, then (IC_L) implies (1) and therefore the FOA is justified. A recent paper by Hwang (2016) proposes a sufficient condition for such global concavity that requires $\Pr(l(\mathbf{x}, e^\dagger) \leq 0 | c^{-1}(z))$ to be convex in z, and shows that his condition is less restrictive than the "convexity of distribution function condition" (CDFC).[15] However, as we will show shortly by an example, Hwang's condition is not satisfied and neither is the global constraint (1) in standard settings. In such cases, the FOA is no longer justified and thus the target effort $e^\dagger$ is not optimal because the agent instead chooses a distant effort lower than $e^\dagger$.[16]

Before turning to the example, we note that the above analysis is relevant only when the first-best effort (denoted e^{FB}) cannot be implemented. For this reason, we make the following assumption on the primitives: [17]

ASSUMPTION 2.

(A4) $\dfrac{\delta}{1-\delta} s(e) \;<\; c(e)$ *for all* $e \geq e^{FB}$.

Observe that the left-hand side of the inequality indicates the maximal bonus that can be paid under the self-enforcement condition (SE). Hence the assumption (A4) tells us that there exists no bonus plan covering the agent's cost from effort for $e \geq e^{FB}$. We define as Ω the set of a pair of contracting parties ω satisfying the assumptions invoked so far:

$$\Omega \;\equiv\; \left\{ \omega = \langle v, c, \overline{\pi}, \overline{u}, \delta \rangle \;\middle|\; \text{satisfying assumptions (A1)} \sim \text{(A4)},\;\; \overline{\pi}, \overline{u} \in \Re_+,\;\; \delta \in (0,1) \right\}.$$

2.1. An Illustrative Example

Consider a unidimensional signal $X \sim N(e, \sigma^2)$, for which we have likelihood ratio $l(x, e) = (x - e)/\sigma^2$. Recall that the FOA bonus plan $\beta^\dagger$ awards the agent a maximal bonus $b^\dagger = \frac{\delta}{1-\delta} s(e^\dagger)$ whenever $l(x, e^\dagger) > 0$. Given this scheme, the agent is paid the bonus $b^\dagger$ with probability

$$\Pr\left(\frac{X - e^\dagger}{\sigma} > 0 \middle| e \right) \;=\; 1 - \Phi\left(\frac{e^\dagger - e}{\sigma} \right),$$

[15] To justify the FOA in formal contracts, in addition to CDFC, the literature often assumes the monotone likelihood ratio property (MLRP) which requires $l(\mathbf{x}, e)$ to be monotone increasing in $\mathbf{x}$ for all e (See Jung and Kim (2015) who propose a distinct set of conditions and show that the MLRP is not necessary for validating the FOA). In our framework, the property is irrelevant to justifying the FOA itself. Instead, it is adopted to guarantee the optimal bonus scheme being monotone in the observed performance.

[16] Kirkegaard (2017a) proposes an alternative approach to justifying FOA in formal contracts. In the environment with risk-neutral parties, however, the proposed condition (Proposition 1) is equivalent to Hwang's condition.

[17] When FOA is valid, it is easily seen that no effort strictly exceeding e^{FB} can be optimal, even if it could be implemented. It appears that this is not generally true when FOA fails, hence we invoke Assumption 2.

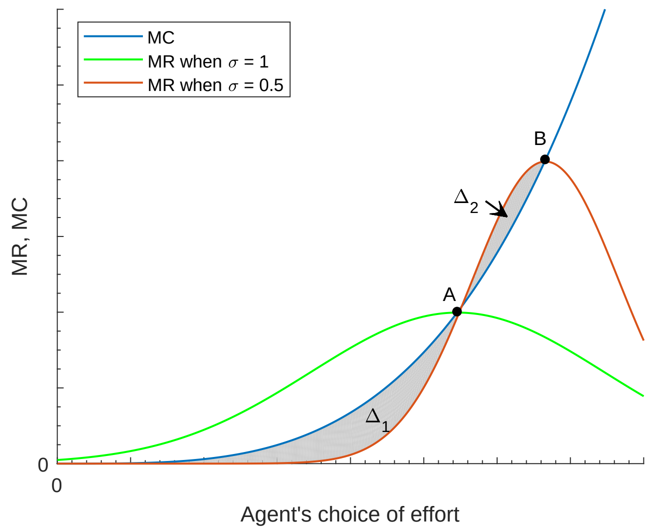


Figure 2: Illustrative example where the first-order approach is not valid.

where $\Phi(\cdot)$ represents the standard normal CDF. Being offered this FOA contract $(w, \beta^\dagger, e^\dagger)$, the agent's marginal (net) gain from exerting effort at e is $b^\dagger\, \Phi'(\frac{e^\dagger - e}{\sigma})\frac{1}{\sigma} - c'(e)$. Since this must be zero at the target effort $e^\dagger$, and since the enforcement constraint (SE) must bind for $b^\dagger$, the target effort is characterized by

$$\frac{\delta}{\sigma(1-\delta)} s(e^\dagger)\Phi'(0) \;=\; c'(e^\dagger). \tag{2}$$

From this condition, it is straightforward that a more informative signal about e (with lower σ) would elevate the agent's marginal revenue from effort (MR, the expression on the left-hand side of (2)), thereby allowing the principal to aim at higher effort. Furthermore, provided that the target effort is below the first best, it in turn allows for a higher bonus pay, so that equilibrium effort and surplus would unambiguously increase. Therefore, a simple comparative static analysis confirms the conventional wisdom that an informative signal alleviates the problem of moral hazard and hence improves on contractual efficiency.

However, there is a caveat to this analysis because the previous local approach is valid only if σ is not too small. Put another way, the FOA fails when the signal becomes very accurate. The reason can be easily seen in Figure 2. When σ is relatively large, the agent's MR from effort (displayed by the green curve) intersects once with the corresponding marginal cost (MC, the blue curve) at point A. Hence at that intersection, the agent's expected payoff is indeed maximized. On the other hand, when σ is small, the agent's MR at the new target effort climbs up to point B, but MR at effort distantly lower than the target is virtually zero, as is depicted by the red curve in Figure 2 (See also Footnote 6). This gives rise to the MR curve intersecting twice with the MC curve. As a consequence, the local approach yields the new target effort at point B, but this effort does not maximize the agent's payoff if the shaded area Δ_2 is smaller than the area Δ_1. If this is the case, the agent would deviate by choosing the minimum level of effort, so the target effort is not implemented.

The example raises two questions: (i) what is an optimal bonus scheme in environments where FOA is not valid, and (ii) does a more informative signal improve on contractual efficiency? We address these questions one by one in the next two sections.

3. Optimal Relational Contracts

In this section, we present a new approach to solving the agency problem in the set Ω. We demonstrate that a hurdle-type bonus scheme like the FOA contract prevails in a wide class of contracting problems, regardless of whether the FOA is valid. Our approach is based on the following condition on the available signal about the agent's hidden action.

DEFINITION 1. *Signal* $\mathbf{X}$ *satisfies the monotone likelihood ratio transformation property* (MLRTP) *if its likelihood ratio* $l(\mathbf{x}, e)$ *possesses the following property: for any* $\kappa \in \Re$ *and* $e, e' \in [0, \bar{e}]$*, there exists a* $\kappa' \in \Re$ *such that*

$$\{\mathbf{x} \in \mathbb{X} | l(\mathbf{x}, e) > \kappa\} \;=\; \{\mathbf{x} \in \mathbb{X} | l(\mathbf{x}, e') > \kappa'\} \tag{3}$$

The condition MLRTP requires that every upper contour set of $l(\mathbf{x}, e)$ can be duplicated by some upper contour set of $l(\mathbf{x}, e')$ with an adjusted level. Analogous to classic consumer theory, the condition endows the likelihood ratios of $\mathbf{X}$ with an *ordinal* property, in the sense that $l(\mathbf{x}', e) \geq l(\mathbf{x}, e)$ for some $(\mathbf{x}', \mathbf{x})$ and some e implies that $l(\mathbf{x}', e') \geq l(\mathbf{x}, e')$ for all $e' \in [0, \bar{e}]$. In other words, for every e and e', $l(\mathbf{x}, e)$ and $l(\mathbf{x}, e')$ are ordinally equivalent. As verified below in Proposition 1, this property is equivalent to either likelihood ratio being a monotone transformation of the other and thus is termed the MLRTP.

Observe that a univariate signal X satisfies the MLRTP if its likelihood ratio $l(x, e)$ is increasing, decreasing or constant in x for all e.[18] See Figure 3-(a). Therefore, the MLRTP subsumes the MLRP as a special case in one-signal cases. In multi-signal cases the two classes overlap, but neither is contained in the other, as is illustrated in the next example.

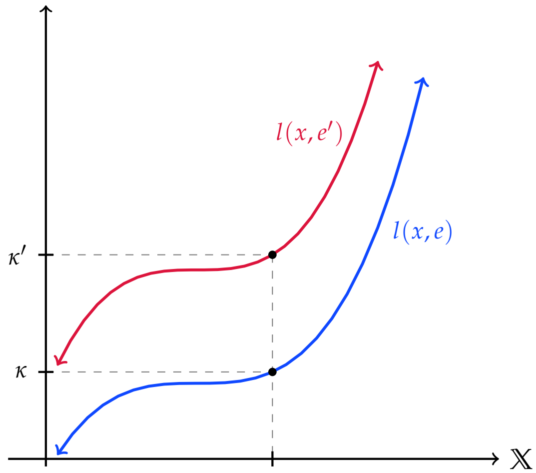
EXAMPLE 1 (MLRTP vs MLRP). *Consider a two-dimensional signal* $\mathbf{X} = (X_1, X_2)$ *defined by CDF*

$$F(\mathbf{x}, e) \;=\; x_1^{\alpha_1(e)} x_2^{\alpha_2(e)}, \quad x_i \in [0,1], \quad \alpha_i(e) > 0 \;\textit{for}\; i = 1, 2.$$

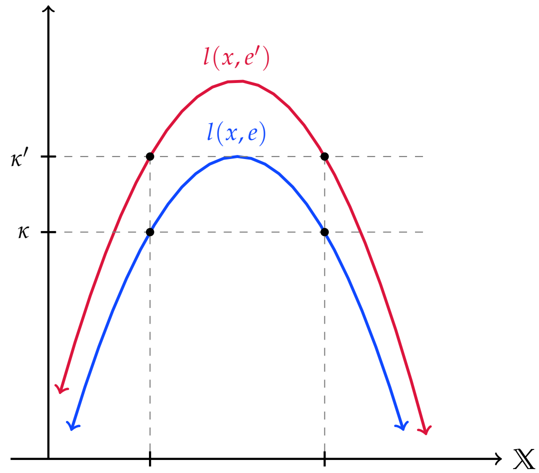
This signal has likelihood ratio

$$l(\mathbf{x}, e) \;=\; \sum_{i=1}^{2} \left(\frac{\alpha_i'(e)}{\alpha_i(e)} + \alpha_i'(e) \log x_i \right),$$

[18] For example, $X \sim N(\mu, \sigma^2)$ with $\sigma = \sigma(e)$ decreasing in e has a likelihood ratio $l(x, e)$ that is hump-shaped in x and yet satisfies the MLRTP. See Figure 3-(b).



(a) Monotone LR



(b) Non-monotone LR

Figure 3: The monotone likelihood ratio transformation property.

where α_i' indicates the derivative of α_i for each $i = 1, 2$. Assume first that $\alpha_1'(e) > 0$ and $\alpha_2'(e) = m\alpha_1'(e)$ with some constant $m \in \Re$ for all admissible e. Then an upper contour set of $l(\mathbf{x}, e)$ takes a form of

$$\{\mathbf{x} \in \mathbb{X} \,|\, l(\mathbf{x}, e) > \kappa\} \;=\; \left\{ \mathbf{x} \in \mathbb{X} \,\middle|\, \log x_1 + m \log x_2 > \left(\kappa - \sum_{i=1}^{2} \frac{\alpha_i'(e)}{\alpha_i(e)} \right) / \alpha_1'(e) \right\}.$$

Since this set can be replicated with the upper contour set of $l(\mathbf{x}, e')$ with an adjusted level κ', the MLRTP *is satisfied for every constant m. On the other hand, for $l(\mathbf{x}, e)$ to be increasing in $\mathbf{x}$ for all e, the constant m must be positive.*

Assume next that $\alpha_1'(e) = 1$ and $\alpha_2'(e) > 0$ for all e, and that there exist e' and e'' such that $\alpha_2'(e') = 1$ and $\alpha_2'(e'') = m_2 \neq 1$. In this case, there is no level $\kappa'' \in \Re$ for which the upper contour set of $l(\mathbf{x}, e'')$ is identical with $\{\mathbf{x} | l(\mathbf{x}, e') > \kappa'\}$. Therefore, MLRP *holds but the* MLRTP *fails for this signal.*

The next result provides a simple characterization for the condition and illustrates its implication.

PROPOSITION 1. *Signal $\mathbf{X}$ satisfies the* MLRTP *if and only if for each e and e', there exists a monotone transformation $\Psi : \Re \to \Re$ satisfying $l(\mathbf{x}, e') = \Psi\left(l(\mathbf{x}, e)\right)$ for all $\mathbf{x} \in \mathbb{X}$. Furthermore, if the condition is met, then for all $\kappa \in \Re$ and $e \in [0, \bar{e}]$*

$$\Pr\big(l(\mathbf{X}, e) > \kappa \,|\, e'\big) \;\text{ is increasing in } e'. \tag{4}$$

PROOF OF PROPOSITION 1: See Appendix A.1. □

The implication (4) of the MLRTP in Proposition 1 has a natural interpretation. It implies that regardless of which effort the principal desires to implement (denoted e), the agent's choice of higher effort (e') becomes more favorable for the ratio $l(\mathbf{X}, e)$ (Milgrom (1981)). Hence the distribution of the ratio can be ranked by first-order stochastic dominance with respect to e'. It is well known that if a univariate signal X satisfies the MLRP, then higher effort is more favorable for the outcome of X, and hence we have $F(x, e'') \leq F(x, e')$ for every $e'' \geq e'$ and x. Since the MLRP forms a bijective association between x and $l(x, e)$, stochastic dominance also holds for the distribution of the likelihood ratio. In case of multivariate signals, the MLRTP ensures (4).

In what follows, we are concerned with signals satisfying the MLRTP. As a leading example, the most natural case $\mathbf{X} = \mu e + \epsilon$, where $\mu = (\mu_1, \cdots, \mu_n)$ and the noise vector $\epsilon = (\epsilon_1, \cdots, \epsilon_n)$ follows a multivariate normal distribution with mean zero vector and covariance matrix $\Sigma = [\sigma_{ij}]$, obeys the MLRTP. Indeed, the likelihood ratio of $\mathbf{X}$ can be written as

$$l(\mathbf{x}, e) \;=\; \sum_{i=1}^{n} m_i (x_i - \mu_i e), \quad \text{where} \;\; m_i \;=\; \sum_{j=1}^{n} \sigma_{ij}^{-1} \mu_j,$$

where σ_{ij}^{-1} are the elements of the inverse matrix Σ^{-1}. The upper contour set $\{\mathbf{x} \in \mathbb{X} | l(\mathbf{x}, e) > \kappa\}$ is thus a half-space of the form

$$\left\{ \mathbf{x} \in \mathbb{X} \;\middle|\; \sum_{i=1}^{n} m_i x_i \;>\; e \sum_{i=1}^{n} m_i \mu_i + \kappa \right\}.$$

It is easy to see that for every effort e', there exists an adjusted level κ' for which the upper contour set of $l(\mathbf{x}, e')$ is identical with the set above.

Another noteworthy example includes the class of signals whose distributions satisfy the spanning condition, i.e. the conditional CDF of outcomes is a convex combination of two CDFs with an effort-dependent weight:[19]

$$F(\mathbf{x}, e) \;=\; \alpha(e) F_1(\mathbf{x}) + (1 - \alpha(e)) F_2(\mathbf{x}), \quad \text{where} \;\; \alpha(e) \in [0, 1] \;\; \text{and} \;\; \alpha'(e) > 0 \quad \forall\, e.$$

Its likelihood ratio at e' is

$$l(\mathbf{x}, e') \;=\; \frac{\partial}{\partial e} \log \left(\alpha(e) \left[\frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} - 1 \right] + 1 \right) \bigg|_{e=e'},$$

where for each $i = 1, 2$, f_i is the density function of F_i. To see that the ratio possesses the ordinal property, we put $\hat{x} \equiv \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} - 1$ to reformulate the ratio as $l(\mathbf{x}, e') = \hat{l}(\hat{x}, e')$. Observe that $\hat{l}(\hat{x}, e')$ is increasing in $\hat{x}$ for every e' as long as $\alpha' > 0$. This monotone property of $\hat{l}$ ensures the existence of an increasing function $m : \Re \to \Re$ satisfying $\hat{l}(m(\kappa), e') = \kappa$ for every constant κ in the range of $l(\mathbf{x}, e')$. Consequently, for another target effort e'', we have $l(\mathbf{x}, e'') = \hat{l}(m(\kappa), e'') = \Psi(\kappa)$ with a monotone transformation Ψ, and hence the MLRTP follows from Proposition 1.

We next investigate the implication of the MLRTP for the form of the optimal bonus scheme. It turns out that the condition enables the principal to evaluate the agent's performance on the basis

[19] Kirkegaard (2017b) develops a method to solve the standard principal-agent problem with multivariate verifiable signals satisfying the spanning condition. In line with this paper, his method does not utilize the FOA.

of the likelihood ratio and thus enables her to focus on the following simple contract.

DEFINITION 2. *A bonus scheme β is a hurdle scheme for the likelihood ratio at target effort $e \in [0, \bar{e}]$ with hurdle $\kappa \in \Re$, if the scheme takes a form of*

$$\beta(\mathbf{x}) = \begin{cases} b\,(>0) & \text{if } l(\mathbf{x}, e) > \kappa \\ 0 & \text{otherwise.} \end{cases}$$

Under this compensation scheme, the agent's performance is evaluated on the basis of a performance index computed from the outcomes $\mathbf{x}$, and the relevant index is the likelihood ratio $l(\mathbf{x}, e)$ at the effort e the principal wants to induce. The scheme is then to reward the agent with a one-step bonus b for all outcomes having index values higher than a hurdle κ. Our first main result can now be stated as follows:

PROPOSITION 2. *Assume the available signal* $\mathbf{X}$ *satisfies the* MLRTP. *Then within class Ω, the optimal bonus scheme is a hurdle scheme for the likelihood ratio at the optimal effort e^*.*

PROOF OF PROPOSITION 2: See Appendix A.2. □

Proposition 2 demonstrates that as long as the available performance evaluation system satisfies the MLRTP, the contract maximizing the joint surplus is a hurdle scheme for all contracting parties in Ω. Hence the principal's problem of how to pay for the agent's performance can be reduced into the problem of choosing a hurdle κ and a bonus b. More importantly, the result does not depend on whether the FOA can be justified or not. Note that the optimal contract derived from the FOA is a special case of the hurdle scheme with $\kappa = 0$. Therefore, Proposition 2 tells us that the optimal contract takes a simple hurdle form in a large class of problems.

To gain insights, recall that for risk-neutral parties, an optimal contract should be designed so as to provide the agent with the strongest incentive to work. Whenever the FOA is valid, equivalently, whenever the local IC constraint (IC_L) is the only relevant one, a way to accomplish this goal is to offer $\beta^\dagger(\mathbf{x})$ that maximizes the *marginal* incentive at the optimal effort e^*:

$$\int_{\mathbb{X}} \beta^\dagger(\mathbf{x})\, l(\mathbf{x}, e^*)\, f(\mathbf{x}, e^*) d\mathbf{x}.$$

This leads to the hurdle scheme for $l(\mathbf{x}, e^*)$ with $\kappa = 0$ being optimal. But as we have discussed in the previous section, this local approach can be justified only if $\beta^\dagger$ satisfies the global IC constraint at the target effort e^*: $u(\beta^\dagger, e) \leq u(\beta^\dagger, e^*)$ for all e. Otherwise, the *total* incentives are not sufficient for inducing e^* from the agent, and thus we have to take into account both incentives when solving the agency problem.

The proof of Proposition 2 proceeds in two steps. We first establish the dominance of a hurdle scheme over any other feasible schemes with regard to the marginal incentive. To be precise, we show that if a non-hurdle scheme β_N implements some effort e^*, then there exists a hurdle scheme

β^* for the likelihood ratio $l(\mathbf{x}, e^*)$ which yields the same expected payoff to the agent as β_N but a higher marginal revenue from effort at e^*. This part of the proof relies on neither FOA nor MLRTP.

Observe that if the hurdle scheme β^* discourages the agent from deviating to lower effort, i.e. if β^* satisfies the global *downward* IC constraints $u(\beta^*, e) \leq u(\beta^*, e^*)$ for all $e \leq e^*$, then the scheme implements higher effort than e^* and thus would Pareto-dominate the non-hurdle scheme β_N. In the second part of the proof, we show that β^* indeed satisfies the constraints if the available signal $\mathbf{X}$ satisfies the MLRTP. As a result, the hurdle scheme is more efficient than others in that it provides the agent with the highest-powered incentives for effort. Therefore, our result suggests that when the FOA bonus scheme $\beta^\dagger$ cannot implement the target effort e^*, the hurdle set at zero should be modified (and the target effort should be modified downwards). This highlights a new role of the hurdle. When FOA is invalid, the adjustment of a hurdle captures a trade-off between on the one hand inducing strong marginal incentives at the target effort, and on the other hand, preventing deviations to distinctly lower effort.

Another meaningful insight on the MLRTP can be found by linking it to a previous condition developed in another contract environment. As Levin (2003) has observed, the stationary relational contract problem is in line with the static problem with two-sided limited liability, in the aspect that both problems impose a lower and upper bound on monetary incentives. In financial contracts between a risk-neutral investor and entrepreneur, Innes (1990) has shown that the additional constraints on liability lead to debt-style contracts being optimal within the class of monotonic contracts. His result has been extended by Poblete and Spulber (2012) to a more general model where the FOA is not necessarily valid. To establish the optimality of debt contracts (in a setting where the slope of the payment scheme is constrained between 0 and 1), Poblete and Spulber (2012) introduce a critical ratio, defined as the marginal return to the principal from increasing the slope of the payment scheme, and assume this ratio to be *regular* in a similar vein as the monotone transformation condition introduced in this paper for the likelihood ratio.[20] With the regularity of critical ratios, they show that the optimal contract exhibits the bang-bang property: it has slope one if the critical ratio exceeds a hurdle but has slope zero otherwise.

When the available signal X is unidimensional, and when the principal's objective $v(e)$ equals $\mathbb{E}[X|e]$, the likelihood ratio plays a similar role to the critical ratio in our setting. To see this, consider a stationary contract (w, β, e) in which the agent's promised utility is fixed at $\overline{u}$. In this case, an increment in bonus pay $\beta(x)$ by Δ over $[x, x + dx]$ would increase the principal's benefit by $\Delta \cdot f_e(x, e)dx$ through the agent's marginal incentive, but at the same time increase the principal's cost by $\Delta \cdot f(x, e)dx$ in order to maintain the continuation payoff. Therefore, the likelihood ratio in relational contracts captures the notion of critical ratios.

Our next result further characterizes the optimal bonus scheme:

PROPOSITION 3. *Suppose signal* $\mathbf{X}$ *satisfies the* MLRTP*. Then the optimal hurdle scheme* $\beta^*(\mathbf{x}) = b^* \mathbb{1}_{\{l(\mathbf{x}, e^*) > \kappa\}}(\mathbf{x})$ *satisfies the following properties:*

[20] In Appendix B, we formally derive the critical ratio and compare their regularity condition with ours in more detail. It turns out that MLRP is sufficient for both conditions, but in general there is no direct connection between them.

(i) the maximal bonus is $b^* = \frac{\delta}{1-\delta} s(e^*)$,

(ii) if the likelihood ratio $l(\mathbf{x}, e)$ decreases (increases) with e for all $\mathbf{x}$, then the optimal hurdle κ is lower (higher) than zero.

PROOF OF PROPOSITION 3: See Appendix A.3 □

The first result of Proposition 3 implies that the self-enforcement condition (SE) must be binding at the optimum, which characterizes the amount of bonus in the hurdle scheme. It now follows that under the MLRTP, an optimal contract can be found by solving for the highest effort $e^* \in [0, e^{FB}]$ that satisfies all downward IC constraints:

$$b^* \Pr(l(\mathbf{X}, e^*) > \kappa | e) - c(e) \ \leq \ b^* \Pr(l(\mathbf{X}, e^*) > \kappa | e^*) - c(e^*), \quad \forall \, e \leq e^*, \tag{IC$_{\text{GD}}$}$$

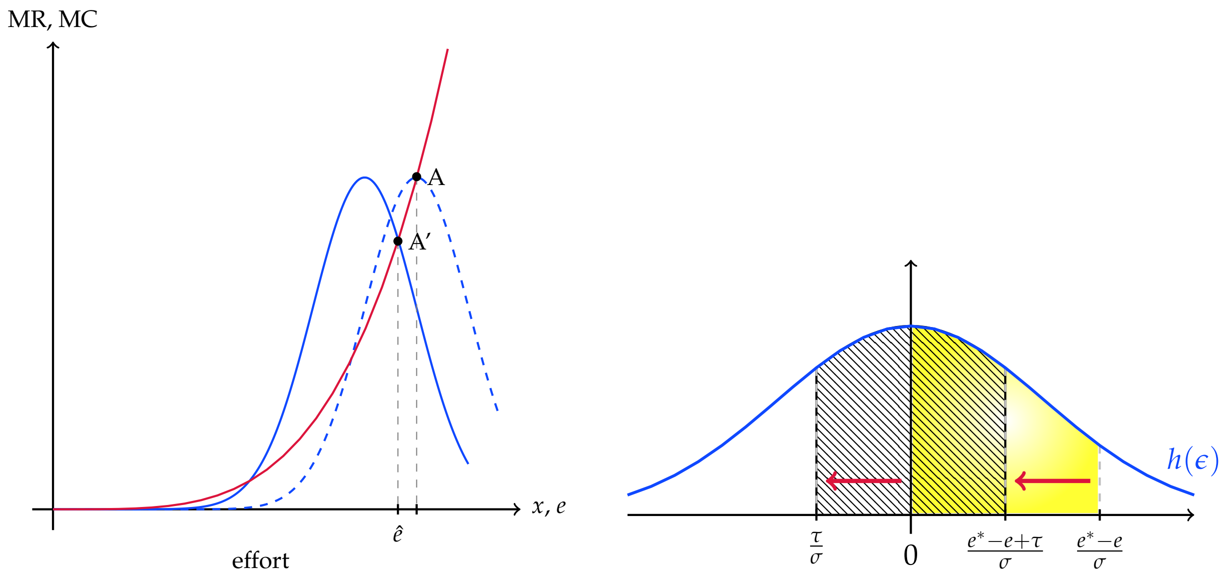
for some hurdle κ and $b^* = \frac{\delta}{1-\delta} s(e^*)$.

The second result illustrates how the optimal hurdle has to be set compared to the FOA contract. It depends on the sign of the derivative $l_e(\mathbf{x}, e)$ whether the optimal contract features a more lenient or strict hurdle than the FOA contract. We show in the proof that if $l(\mathbf{x}, e)$ decreases with e, then the probability of getting a bonus $\Pr(l(\mathbf{X}, e^*) > \kappa | e)$ satisfies decreasing differences in $(\kappa; e)$ as long as the hurdle κ remains positive. Hence the agent's expected payoff function becomes submodular in $(\kappa; e)$, which in turn implies that the principal's and agent's actions are strategic substitutes. In this case, a more lenient hurdle induces higher effort. In contrast, an increasing likelihood ratio with e leads to strategic complementarity between κ and e.

Some intuition for a negative hurdle can be obtained from the example in the previous section. In Figure 4-(a), the blue and red curves depict the agent's marginal gain and marginal cost from effort, respectively, for the case of a signal $X \sim N(e, \sigma^2)$, where the bonus hurdle has been set at $\kappa = 0$ in accordance with the FOA contract.[21] In the case depicted, the signal variance is small, and the FOA solution for effort (given by the intersection point A where the dotted curve achieves the maximum value) is a local but not a global optimum for the agent. Here ceteris paribus, a variation of the hurdle κ will entail a horizontal shift of the marginal revenue curve. The blue curve corresponds to some negative hurdle $\kappa < 0$ for the likelihood ratio. This more lenient hurdle undermines the agent's marginal incentives for "high" effort but instead strengthens his total incentives for such effort. Effort $\hat{e}$ at the highest intersection (point A') of the marginal revenue curve and the marginal cost curve is now a global optimum for the agent.

The next example illustrates how to utilize the results in this section by characterizing the optimal contract in a simple setting.

[21] This bonus hurdle for the likelihood ratio corresponds to a hurdle $x > e^\dagger$ for the signal outcome x, and the marginal revenue is then proportional to the normal density $\Phi'(\frac{e^\dagger - e}{\sigma})$.



(a) Marginal Gains

(b) Total Gains

Figure 4: The effect of lowering hurdle κ on the agent's marginal and total gains from effort

3.1. Example: Characterization of Optimal Bonus Scheme

Consider a unidimensional noisy signal of effort $X = e + \epsilon\sigma$, where ϵ has a log-concave density $h(\cdot)$ with a unique mode at zero. The likelihood ratio of X at e can be written as

$$l(x,e) \;=\; -\frac{1}{\sigma}h'\left(\frac{x-e}{\sigma}\right)/h\left(\frac{x-e}{\sigma}\right),$$

which is increasing in x but decreasing in e. In light of Proposition 2 and 3-(i), we can restrict attention to a hurdle scheme $\beta(x) = b^*\mathbb{1}_{\{l(x,e^*)>\kappa\}}(x)$ with $b^* = \frac{\delta}{1-\delta}s(e^*)$ for optimal contracts. For each hurdle κ, there exists a unique threshold τ such that $l(x,e^*) > \kappa$ iff $x - e^* > \tau$. Hence we can express the probability of getting a bonus as

$$1 - \Pr(l(X,e^*) \leq \kappa|e) \;=\; \Pr\left(X - e^* > \tau|e\right) \;=\; 1 - H\left(\frac{e^*-e+\tau}{\sigma}\right),$$

where $H(\cdot)$ indicates the CDF of ϵ.

If the FOA is valid, the optimal hurdle is determined by the first-order condition of the agent's problem:

$$b^*h\left(\frac{\tau}{\sigma}\right) \;=\; \sigma c'(e^*). \tag{5}$$

In order to induce the highest effort e^* under this local constraint, the principal must set $\tau = 0$, or equivalently $\kappa = 0$ in the hurdle scheme, as the density function h achieves the maximum at zero. This contract is indeed optimal if it satisfies the global downward constraints: that is, $\tau = 0$ satisfies the following set of inequalities:

$$H\left(\frac{e^*-e+\tau}{\sigma}\right) - H\left(\frac{\tau}{\sigma}\right) \geq \frac{c(e^*)-c(e)}{b^*} \qquad \forall\, e \leq e^*. \tag{6}$$

If the inequality does not hold for some e, put differently, if the highlighted are in Figure 4-(b), $H((e^*-e)/\sigma) - H(0)$, is smaller than $\frac{c(e^*)-c(e)}{b^*}$ for some $e \leq e^*$, then the scheme with $\tau = 0$ does not induce the agent to choose the desired effort e^*. Proposition 3-(ii) tells us that the way to resolve this incentive problem is to lower the hurdle. By setting $\tau < 0$, the principal can adequately increase the net gain from exerting e^* to the agent as is displayed in Figure 4-(b). The example highlights a key trade-off in determination of the hurdle. By lowering the hurdle from $\tau = 0$, the agent's downward incentive constraints (6) will be relaxed, but the agent's marginal incentives for "high" effort (5) will be undermined. The optimal contract must find the right balance between these two effects.

When the FOA is not justified, the optimal contract can be derived as follows. At an optimal non-zero threshold τ, some downward IC constraint must be binding (for otherwise the FOA is valid), and the corresponding effort, say $e_0 < e^*$, must be a local optimum for the agent's problem. Hence we have

$$H\left(\frac{e^*-e_0+\tau}{\sigma}\right) - H\left(\frac{\tau}{\sigma}\right) = \frac{c(e^*)-c(e_0)}{b^*} \tag{7}$$

and

$$b^*h\left(\frac{e^*-e_0+\tau}{\sigma}\right) \leq \sigma c'(e_0), \qquad e_0 \geq 0, \tag{8}$$

where the last two inequalities hold with complementary slackness at the local optimum e_0. In addition, the optimal effort e^* must be a local (and interior) optimum for the agent, and the self-enforcement condition must be binding (Proposition 3-(i)). This gives us,

$$b^*h\left(\frac{\tau}{\sigma}\right) = \sigma c'(e^*) \quad \text{and} \quad b^* = \frac{\delta}{1-\delta}s(e^*). \tag{9}$$

These are necessary conditions. If in addition we know that the agent's payoff has at most two local maxima (as is the case when ϵ is normal and $c'(e)$ is linear), the conditions (7, 8, 9) will also be sufficient to determine τ, e^* and e_0.[22]

4. Value of Information

In the previous section, we studied the properties of an optimal bonus scheme in the stationary environment for a *given* signal. If the signal satisfies a mild condition, then the optimal scheme takes a simple hurdle form for all contracting parties in the class Ω. Utilizing this fact, we now turn to the problem of choosing a signal, i.e., a performance measurement system. To state our problem, suppose that there are two available non-verifiable signals satisfying the MLRTP, say $\mathbf{X}$

[22] In standard agency models with a univariate signal, Ke and Ryan (2018) show that if the signal satisfies the MLRP, then there exist at most two local maxima in the agent's problem. Thus, if the FOA is not valid, introducing one additional constraint (7) which they call a *no-jump* constraint is enough for solving the problem. In our example, there may be multiple solutions for e_0, suggesting that the downward constraints may bind at several effort levels below e^*.

with support $\mathbb{X} \subset \Re^n$ and $\mathbf{Y}$ with support $\mathbb{Y} \subset \Re^m$. Our objective here is to propose a new criterion under which the principal prefers to measure the agent's performance and write a contract based on signal $\mathbf{X}$ rather than signal $\mathbf{Y}$.

There are a few previous works investigating the nature of a more informative signal in a principal-agent framework. However, most attention has been devoted to explicit (or formal) contracts, that is, to models of contracting with a verifiable signal and risk-averse agent. In such contracts, the agency costs arises from the tradeoff between incentive provision and risk. As a result, the incentives are constrained: for example, selling the project to the agent is never optimal. The literature has developed criteria for a signal to be more informative and thus better alleviate agency costs, which include the informativeness criterion by Holmström (1979) and the mean-preserving spread (MPS) criterion by Kim (1995), among others. In a relational contract with a risk-neutral agent, on the other hand, it is the enforcement problem rather than the agent's risk attitude that constrains the feasible set of contracts and thus hinders a contract from implementing the first-best. The different source of the agency cost suggests that a direct application of the existing criteria to non-verifiable signals is inappropriate.

In this section, we establish a new criterion for a more informative signal tailored to relational contracts. In general, a signal $\mathbf{X}$ is more informative than another $\mathbf{Y}$ if writing a contract based on $\mathbf{X}$ is more effective in reducing the agency costs than doing so based on $\mathbf{Y}$. In our framework, a more informative signal enables the principal to implement higher effort and thus both parties to achieve a higher surplus in the optimal contract.[23] We attempt to present a robust condition with respect to the characteristics of the model, under which signal $\mathbf{X}$ induces higher effort than signal $\mathbf{Y}$ for any pair of parties in the class Ω. To formalize this idea, select one agency problem ω from Ω, and let $e_{\mathbf{X}}(\omega)$ and $e_{\mathbf{Y}}(\omega)$ denote the optimal effort, i.e., the highest effort implemented by a contract based on signal $\mathbf{X}$ and $\mathbf{Y}$, respectively.

DEFINITION 3. *Signal $\mathbf{X}$ is more informative than signal $\mathbf{Y}$ within class Ω if $e_{\mathbf{X}}(\omega) \geq e_{\mathbf{Y}}(\omega)$ for all $\omega \in \Omega$.*

One notable feature of this notion is that the principal's objective v is not directly affected by her choice of signals but only indirectly affected through the agent's choice of effort. Whether the principal designs a contract with signal $\mathbf{X}$ or $\mathbf{Y}$, her expected return is determined by the agent's productive inputs, not by the choice of a performance measure.[24] Abstracting away this direct effect, we focus on the problem of choosing an ideal performance measure that strengthens incentives in relational contracts.

We present a statistical criterion that characterizes a more informative signal. The results established in the previous section suggest that the criterion pertains to the likelihood ratio of a signal.

[23] Since the way to distribute the surplus does not influence the agent's incentives (due to the base salary, see Theorem 1 in Levin (2003)), signal $\mathbf{X}$ is more informative than signal $\mathbf{Y}$ if a contract of the form $\beta(\mathbf{X})$ Pareto-improves any contracts of the form $\beta(\mathbf{Y})$.

[24] This assumption can be easily justified in two cases: (1) the realized returns are not observed (or very hard to be assessed) by the parties at the stage of payment and thus are not part of $\mathbf{x}$ or $\mathbf{y}$, or (2) the returns are observable and determined by part of $\mathbf{x}$ or $\mathbf{y}$ (for instance, the first element of each signal), but both signals have the same marginal distribution on that part.

Let $\mathbf{X} \sim F(\cdot, e)$ and $\mathbf{Y} \sim G(\cdot, e)$, where F and G are the respective CDFs, and let $f(\mathbf{x}, e)$ and $g(\mathbf{y}, e)$ denote the respective densities of each signal. Given an effort e^*, which we shall call the target effort hereafter, we denote the CDF of the likelihood ratio $l(\mathbf{x}, e^*)$ conditional on the agent's choice of effort e by

$$L_{\mathbf{X}}(\kappa, e) \;\equiv\; \Pr\Big(l(\mathbf{X}, e^*) \leq \kappa \,\Big|\, e\Big) \;=\; \int_{\mathbb{X}} \mathbb{1}_{\{l(\mathbf{x}, e^*) \leq \kappa\}}(\mathbf{x}) f(\mathbf{x}, e) d\mathbf{x}.$$

The distribution $L_{\mathbf{Y}}(\kappa, e) \equiv \Pr\left(l(\mathbf{Y}, e^*) \leq \kappa \,|\, e\right)$ can be defined in a similar way. Note that under the MLRTP, the distribution of $l(\mathbf{x}, e')$ at another target effort $e' \neq e^*$ is isomorphic to the distribution of $l(\mathbf{x}, e^*)$ with an adjusted κ', that is, $\Pr(l(\mathbf{X}, e') < \kappa | e) = L_{\mathbf{X}}(\kappa', e)$. Hence the ordinal property of the MLRTP allows us to drop the target effort in the distribution of the ratio.

With this notation, our informativeness criterion and main result of this section can be stated as follows:

DEFINITION 4. *Signal* $\mathbf{X}$ *dominates signal* $\mathbf{Y}$ *in the likelihood ratio order if for every* $\kappa \in \Re$ *and target effort* $e^* \in [0, \bar{e}]$, *there exists a* $\kappa' \in \Re$ *satisfying*

$$L_{\mathbf{Y}}(\kappa, e) - L_{\mathbf{Y}}(\kappa, e^*) \;\leq\; L_{\mathbf{X}}(\kappa', e) - L_{\mathbf{X}}(\kappa', e^*) \quad \text{for all} \;\; e < e^*. \tag{L}$$

If (L) *holds between the two signals, we write* $\mathbf{X} \succ_{\mathrm{L}} \mathbf{Y}$.

PROPOSITION 4. *Suppose that two signals* $\mathbf{X}$ *and* $\mathbf{Y}$ *satisfy the* MLRTP. *Then* $\mathbf{X}$ *is more informative than* $\mathbf{Y}$ *within class* Ω *if and only if* $\mathbf{X} \succ_{\mathrm{L}} \mathbf{Y}$.

PROOF OF PROPOSITION 4: See Appendix A.4. □

Roughly speaking, the statistical order defined in (L) compares the variability of the likelihood ratios in response to the agent's choice of effort. The difference on the left-hand side of (L) indicates the change in the distribution of $l(\mathbf{Y}, e^*)$ when the agent does not follow the instruction e^* but deviates to some lower effort $e < e^*$. With this in hand, the difference can be interpreted as the amount of information regarding the agent's possible deviations conveyed by the informational variable $l(\mathbf{Y}, e^*)$, or simply signal $\mathbf{Y}$. Hence the likelihood ratio order $\mathbf{X} \succ_{\mathrm{L}} \mathbf{Y}$ implies that the maximal amount of information contained in $\mathbf{Y}$ is outweighed by that contained in $\mathbf{X}$, regardless of the desired effort e^*. Consequently, the principal can more effectively control the agent's hidden action by designing a contract based on signal $\mathbf{X}$ rather than $\mathbf{Y}$.

To gain more insights behind Proposition 4, suppose that the target effort e^*, bonus b^* and hurdle κ constitute an optimal contract under signal $\mathbf{Y}$, and consider the downward IC constraints ($\mathrm{IC}_{\mathrm{GD}}$), which now can be written as

$$b^* \left(L_{\mathbf{Y}}(\kappa, e) - L_{\mathbf{Y}}(\kappa, e^*)\right) \;\geq\; c(e^*) - c(e), \quad \forall\, e \leq e^*.$$

Note that $1 - L_{\mathbf{Y}}(\kappa, e)$ indicates the probability that the agent's performance will surpass the hurdle κ and thus the agent receive the bonus b^* with choice of effort e. The left-hand side of the inequality above is thus the expected loss of bonus when the agent deviates to a lower effort e. The inequality states that in order to implement e^*, this loss should exceed the effort costs saved by any possible downward deviations.

If signal $\mathbf{X}$ dominates $\mathbf{Y}$ in the likelihood ratio order, then there exists a hurdle scheme with an adjusted κ' for signal $\mathbf{X}$ such that a deviation to lower effort $e < e^*$ entails a larger reduction in the probability to pass the relevant hurdle under $\mathbf{X}$ than under $\mathbf{Y}$. Given the same bonus b^*, any deviations are then even less attractive under signal $\mathbf{X}$, and this guarantees that the target effort e^* can be implemented under this signal. In this sense, the likelihood ratio order is sufficient for $\mathbf{X}$ to be more informative than $\mathbf{Y}$. Furthermore, Proposition 4 tells us that for $\mathbf{X}$ to be more informative for any pair of parties in Ω, the condition (L) is also necessary. To be specific, if the two signals cannot be ranked by the likelihood ratio order, then there exists a pair of contracting parties in Ω for whom it is more efficient to evaluate the agent's performance based on $\mathbf{Y}$ rather than $\mathbf{X}$. Consequently, $\mathbf{X} \succ_L \mathbf{Y}$ presents the way to order non-verifiable signals in relational contracts.

To see how the likelihood ratio order is related to other criteria in the literature, consider the following notion of ordering univariate signals, first introduced by Lehmann in the field of statistical decision theory.[25]

DEFINITION 5. (LEHMANN (1988)) *A univariate signal X is more precise about unknown parameter $e \in [0, \bar{e}]$ than another univariate signal Y if for every realization y of Y, there exists an increasing function $T_y : [0, \bar{e}] \to \mathbb{X}$ such that*

$$F(T_y(e), e) \;=\; G(y, e) \quad \text{for all } e. \tag{P}$$

If (P) holds between the two signals, we write $X \succ_P Y$.

It is the monotonicity of T_y that is essential for X to be statistically more precise than Y in Lehmann's notion. To see its role, consider two uniform distributions $X \sim U[e - \sigma/2, \; e + \sigma/2]$ and $Y \sim U[e - 1/2, \; e + 1/2]$. Equating their CDFs, we can compute the associated T-transformation:

$$T_y(e) \;=\; \sigma(y - e) + e \quad \text{for } y \in \left[e - \frac{1}{2}, \, e + \frac{1}{2}\right].$$

It is straightforward to see that the obtained T_y is increasing in e if and only if $\sigma < 1$, i.e., if the density of X is more clustered around e. For a signal that is exposed to an additive shock, Lehmann's order provides a more intuitive and complete ranking than Blackwell's sufficiency.[26]

With this order we have the following result:

PROPOSITION 5.

[25] For univariate signals X and Y satisfying the MLRP, Lehmann (1988) proved that X is more informative than Y in a statistical decision problem (with some restrictions on the decision maker's underlying payoff function) if and only if $X \succ_P Y$. Comparing with the statistical order based on sufficiency developed by Blackwell (1951, 1953), (P) is a more complete and intuitive order, and moreover, is easier to check so that it has been applied to several economic environments since Persico (2000).

[26] The given example is due to Lehmann (1988), who showed that X is sufficient for Y only if $\sigma = 1/k$ where $k \geq 1$ is a natural number.

(i) *Suppose that $l(\mathbf{X}, e^*)$ is more precise than $l(\mathbf{Y}, e^*)$ for all $e^* \in [0, \bar{e}]$. That is, for every $\kappa \in \Re$ in the support of $l(\mathbf{Y}, e^*)$, there exists an increasing function $T_\kappa(e)$ such that*

$$L_{\mathbf{X}}\Big(T_\kappa(e),\, e\Big) \;=\; L_{\mathbf{Y}}(\kappa,\, e) \quad \forall\, e.$$

Then $\mathbf{X} \succ_L \mathbf{Y}$.

(ii) *For a comparison of univariate signals satisfying the strict* MLRP, *signal X is more precise than signal Y if and only if the likelihood ratio $l(X, e^*)$ of X is more precise than the ratio $l(Y, e^*)$ of Y for every e^*. Consequently,*

$$X \succ_{\mathrm{P}} Y \quad \textit{implies} \quad X \succ_{\mathrm{L}} Y.$$

PROOF OF PROPOSITION 5: See Appendix A.5. □

Proposition 5 has two implications. First, a signal with a more precise likelihood ratio dominates in the likelihood ratio order, and hence allows higher effort to be implemented in a relational contract. As the principal uses the likelihood ratio as a key indicator to decide whether to pay a bonus in the optimal contract, she prefers to evaluate the agent's performance with a more precise likelihood ratio. Second, for univariate signals satisfying the strict MLRP, the notion of a more precise likelihood ratio is identical with Lehmann's original notion of a more precise signal. To gain some intuition for their equivalence, observe that the strict MLRP gives rise to a bijective mapping between the distribution of a signal and the distribution of its likelihood ratio. Consequently, the first implication leads to a simple and intuitive way to order non-verifiable signals: a more precise signal is more informative in relational contracts.

The next example illustrates how Proposition 5 can be applied to the previous example in Section 3.

EXAMPLE 2. *Let $X = e + \epsilon\sigma_1$ and $Y = e + \epsilon\sigma_2$, where the additive noise ϵ has the CDF $H(\cdot)$ as in the example in Section 3.1. We can write the distribution of each signal as $F(x, e) = H\left(\frac{x-e}{\sigma_1}\right)$ and $G(y, e) = H\left(\frac{y-e}{\sigma_2}\right)$. The associated T-transformation is therefore*

$$T_y(e) \;=\; \frac{\sigma_1}{\sigma_2} \cdot y + \frac{\sigma_2 - \sigma_1}{\sigma_2} \cdot e,$$

which is increasing in e whenever $\sigma_2 > \sigma_1$. Consequently, X is more precise than Y, and hence more informative than Y if $\sigma_1 < \sigma_2$. The result holds regardless of whether FOA can be justified or not.

As seen in the example above, Proposition 5 provides a simple method to check which signal is more informative, without a further computation work for the likelihood ratio and its distribution. The converse of (ii) in Proposition 5 is generally not true, but we have the following result:

COROLLARY 1. *Given two univariate signals X and Y satisfying the strict* MLRP, *suppose that the associated T-transformation between X and Y is additively separable. Then $X \succ_{\mathrm{L}} Y$ if and only if $X \succ_{\mathrm{P}} Y$.*

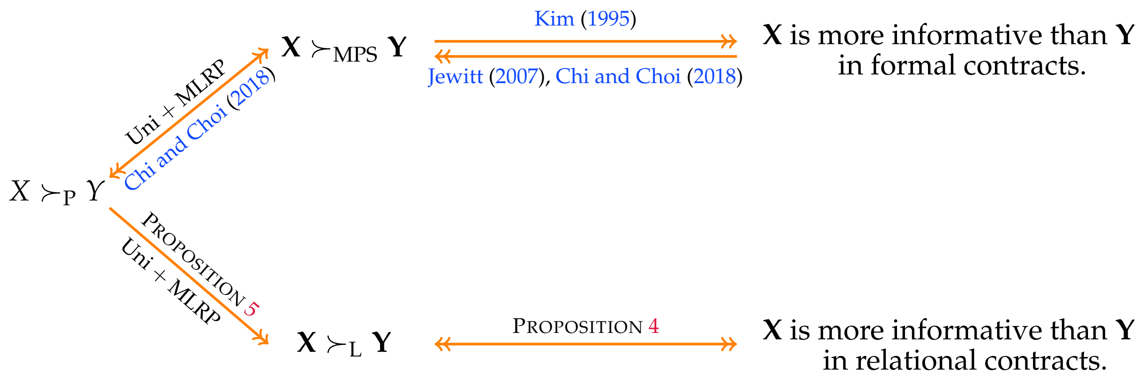


Figure 5: Link to Alternative Signal Orderings

If the T-transformation is additively separable, then $X \succ_{\text{L}} Y$ implies $X \succ_{\text{P}} Y$ so that the two stochastic orders are equivalent. To see this, observe that in case of univariate signals satisfying the MLRP, the likelihood ratio order (L) can be written in the following fashion: for every $y \in \mathbb{Y}$, there exists a $x \in \mathbb{X}$ such that $G_e(y, e) \geq F_e(x, e)$. Taking the derivative of both sides of the identity $F(T_y(e), e) = G(y, e)$ with respect to e, we have

$$G_e(y, e) \;=\; f(T_y(e), e) \cdot \frac{\partial T_y(e)}{\partial e} \;+\; F_e(T_y(e), e). \tag{10}$$

Since $T_y(e)$ is a bijective function of y for each e, for the existence of x satisfying $G_e(y, e) \geq F_e(x, e)$ for all y, $\partial T_y(e)/\partial e$ must be nonnegative at least for some y. Hence if T is additively separable, we have a nonnegative derivative of T_y for every y, leading to $X \succ_{\text{P}} Y$. For signals generated by an additive shock as in the previous example, the T-transformation is indeed additively separable, and consequently X dominates Y in the likelihood ratio order if and only if $\sigma_2 > \sigma_1$.

Proposition 5 also provides a link to the MPS criterion proposed by Kim (1995) in explicit contracts.[27] Just as the likelihood ratio order in relational contracts, the MPS criterion serves as an informativeness criterion for ranking verifiable signals in explicit contracts (but in the latter case under the assumption that FOA is justified).[28] The likelihood ratio order is in line with the MPS criterion in the aspect that both criteria are based on the variability of the likelihood ratio. In particular, for a comparison of univariate signals satisfying the MLRP, these two criteria are connected through Lehmann's order. Recently, Chi and Choi (2018) prove that within the class of such signals, the MPS criterion is equivalent to Lehmann's order. Hence it follows from Proposition 5-(ii) that $X \succ_{\text{MPS}} Y$ implies $X \succ_{\text{L}} Y$. Figure 5 displays this linkage. However, the two criteria do not mutually imply each other in general, since they adopt different notions of variability.[29] This distinction brings forth a new approach to ranking signals in relational contracts,

[27] With our notation, Kim's criterion $\mathbf{X} \succ_{\text{MPS}} \mathbf{Y}$ can be stated as follows: the distribution $L_{\mathbf{X}}(\kappa, e) \equiv \Pr(l(\mathbf{X}, e) \leq \kappa | e)$ is more dispersed than the distribution $L_{\mathbf{Y}}(\kappa, e) \equiv \Pr(l(\mathbf{Y}, e) \leq \kappa | e)$ for all e in the sense of a mean-preserving spread.

[28] Kim (1995) establishes the sufficiency part of the MPS criterion for informativeness. Jewitt (2007) and more recently Chi and Choi (2018) establish the necessity of the criterion by using the dual approach.

[29] Another notable difference is that the MPS criterion hinges upon FOA so that it compares the variability of the

enlightening the different source of agency costs in the two types of contracts.

5. Conclusion

Performance measurement and design of incentive schemes are central issues in agency theory. The main purpose of this paper has been to develop a new criterion to characterize a better measurement system in relational contracts. Relational incentive contracts with non-verifiable information are subject to the enforcement problem, which is the main source of agency costs when the contracting parties are risk neutral. This results in a distinctive criterion, compared to the existing informativeness criteria for verifiable signals in formal contracts with risk averse agents. For its application to a wide class of signals, we also developed a condition, the MLRTP, under which the optimal bonus scheme takes a simple hurdle form, regardless of whether the FOA is applicable. In such settings, the likelihood ratio order serves as a criterion for the efficient measurement system.

A limitation of our model is that it is confined to one-dimensional effort. Many agents are involved in multi-tasking, and the issues analyzed here are of course also relevant for such settings. When the FOA is valid in a multi-task problem, the likelihood ratios on the agent's various tasks will play a key role in the optimal bonus scheme for a relational contract (Kvaløy and Olsen (2017)). It would be interesting to see if the results developed in this paper can be extended into a multi-task setting.

While this paper is confined to relational contracting with only non-verifiable performance measurements, in reality there is often a combination of verifiable and non-verifiable measures available. A recent paper by Miller, Olsen and Watson (2018) develops a general framework to analyze relational contracting in such settings, where contracts will have self-enforced as well as externally enforced elements. Extending our results regarding optimal bonus schemes and ranking of performance measurement systems to such environments would also be interesting and useful.

A. Omitted Proofs

A.1. Proof of Proposition 1

We first establish the equivalence between (i) the MLRTP and (ii) existence of a monotone transformation Ψ satisfying $l(\mathbf{x}, e') = \Psi(l(\mathbf{x}, e))$. The fact that (ii) implies (i) can be easily shown by setting $\kappa' = \Psi^{-1}(\kappa)$. To prove the other direction, suppose that the set of likelihood ratios $\{l(\cdot, e),\ e \in [0, \bar{e}]\}$ satisfies the MLRTP. Given a choice of effort e, we can make a partition of the image of $l(\mathbf{x}, e)$ with an increasing sequence $\{\kappa_n\}_{n=0}^{N}$:

$$\inf_{\mathbf{x}\in\mathbb{X}} l(\mathbf{x}, e) = \kappa_0 < \kappa_1 < \cdots < \kappa_{N-1} < \kappa_N = \sup_{\mathbf{x}\in\mathbb{X}} l(\mathbf{x}, e).$$

likelihood ratios in response to the agent's *local* deviation from each target effort, whereas our criterion compares the variability in response to *global* downward deviations.

Observe that the increasing sequence $\{\kappa_n\}$ yields a simple function converging to $l(\mathbf{x}, e)$ pointwise:

$$\sum_{n=1}^{N} \kappa_n \mathbb{1}_{\{\kappa_{n-1} < l(\mathbf{x},e) \le \kappa_n\}}(\mathbf{x}) \quad \xrightarrow{N\uparrow\infty} \quad l(\mathbf{x}, e).$$

To the given sequence $\{\kappa_n\}$, we construct a corresponding sequence $\{\kappa'_n\}$ such that $\{\mathbf{x} \in \mathbb{X} \,|\, l(\mathbf{x}, e) > \kappa_n\} = \{\mathbf{x} \in \mathbb{X} \,|\, l(\mathbf{x}, e') > \kappa'_n\}$ for each n. Since the collection of upper contour sets $\langle\{\mathbf{x}|l(\mathbf{x}, e) > \kappa_n\}\rangle_{n=0}^{N}$ is descending in the sense that $\{\mathbf{x} \,|\, l(\mathbf{x}, e) > \kappa_n\} \supseteq \{\mathbf{x} \,|\, l(\mathbf{x}, e) > \kappa_{n+1}\}$ for each n, the constructed sequence $\{\kappa'_n\}$ must be increasing. This allows us to define a real-valued function Ψ that transforms κ_n to κ'_n, i.e., $\Psi(\kappa_n) = \kappa'_n$. It is self-evident that Ψ is an order-preserving transformation. Using the MLRTP and the transformation Ψ, we can rewrite the simple function generated by $\{\kappa'_n\}$ into

$$\sum_{n=1}^{N} \kappa'_n \mathbb{1}_{\{\kappa'_{n-1} < l(\mathbf{x},e') \le \kappa'_n\}}(\mathbf{x}) \;=\; \sum_{n=1}^{N} \Psi(\kappa_n) \mathbb{1}_{\{\kappa_{n-1} < l(\mathbf{x},e) \le \kappa_n\}}(\mathbf{x}). \tag{11}$$

The desired result then follows from the standard approximation argument by which the expression on each side of (11) converges to $l(\mathbf{x}, e')$ and $\Psi(l(\mathbf{x}, e))$, respectively, as N grows large.

For the proof of the stochastic dominance, we take the derivative of $\Pr(l(\mathbf{X}, e) > \kappa|e')$ with respect to e' to get

$$\frac{\partial}{\partial e'}\Pr\big(l(\mathbf{X}, e) > \kappa|e'\big) \;=\; \mathbb{E}\left[l(\mathbf{X}, e')\mathbb{1}_{\{l(\mathbf{X},e)>\kappa\}}(\mathbf{X}) \,|\, e'\right] \;=\; \mathbb{E}\left[l(\mathbf{X}, e')\mathbb{1}_{\{l(\mathbf{X},e')>\kappa'\}}(\mathbf{X}) \,|\, e'\right] \;\ge\; 0,$$

where the second equality is due to the MLRTP and the last inequality is due to the fact that the expectation of $l(\mathbf{X}, e')$ is zero for all e'.

A.2. Proof of Proposition 2

Given a bonus scheme $\beta : \mathbb{X} \to \Re$ and the agent's choice of effort $e \in [0, \bar{e}]$, define the agent's expected payoff (less the base salary) as

$$u(\beta, e) \;=\; \int_{\mathbb{X}} \beta(\mathbf{x})\, f(\mathbf{x}, e) d\mathbf{x} - c(e) \;=\; \mathbb{E}[\beta(\mathbf{X})|e] - c(e).$$

Also, we define by $u_e(\beta, e) \equiv \partial u(\beta, e)/\partial e$ the marginal incentive at e given β.

The next lemma is a simple application of the integral inequalities (Theorem 3) in Banks (1963), which plays a key role in the proof of Proposition 2. The lemma shows that for provision of the strongest marginal incentives at target effort e^o, a hurdle scheme is more effective than any other bonus schemes. Note that the lemma does not rely on the MLRTP.

LEMMA 1. *Let β_N be an arbitrary bonus scheme satisfying $0 \le \beta_N(\mathbf{x}) \le b$ for all $\mathbf{x} \in \mathbb{X}$. Then for each admissible effort e^o, there exists a hurdle scheme β with maximal bonus b such that*

$$u(\beta, e^o) \;=\; u(\beta_N, e^o) \quad \text{and} \quad u_e(\beta, e^o) \;\ge\; u_e(\beta_N, e^o).$$

PROOF OF LEMMA 1: Given a level of effort $e^o \in [0, \bar{e}]$ and the upper bound b for β_N, construct a hurdle scheme for likelihood ratio $l(\mathbf{x}, e^o)$ with hurdle $\kappa \in \Re$ as follows:

$$\beta(\mathbf{x}) = \begin{cases} 0 & \text{if } l(\mathbf{x}, e^o) < \kappa \\ b & \text{otherwise,} \end{cases}$$

where the hurdle κ is set such that the two schemes yield the same expected payoff at e^o. The intermediate value theorem guarantees the existence of such κ.

With this scheme in hand, observe that the difference of marginal incentives between the two schemes at $e = e^o$ is

$$\begin{aligned} u_e(\beta, e^o) - u_e(\beta_N, e^o) &= \int_{\mathbb{X}} \Big(\beta(\mathbf{x}) - \beta_N(\mathbf{x})\Big) l(\mathbf{x}, e^o) f(\mathbf{x}, e^o) d\mathbf{x} \\ &= -\int_{l(\mathbf{x}, e^o) < \kappa} \beta_N(\mathbf{x}) l(\mathbf{x}, e^o) f(\mathbf{x}, e^o) d\mathbf{x} \\ &\qquad\qquad + \int_{l(\mathbf{x}, e^o) \geq \kappa} \Big(b - \beta_N(\mathbf{x})\Big) l(\mathbf{x}, e^o) f(\mathbf{x}, e^o) d\mathbf{x} \\ &\geq \kappa \left\{ -\int_{l(\mathbf{x}, e^o) < \kappa} \beta_N(\mathbf{x}) f(\mathbf{x}, e^o) d\mathbf{x} + \int_{l(\mathbf{x}, e^o) \geq \kappa} \Big(b - \beta_N(\mathbf{x})\Big) f(\mathbf{x}, e^o) d\mathbf{x} \right\}, \end{aligned}$$

where the last inequality takes a strict form whenever the set $\{\mathbf{x} \in \mathbb{X} \mid \beta_N(\mathbf{x}) \neq \beta(\mathbf{x})\}$ has positive measure. Since the expression in the curly bracket on the bottom line above boils down to $u(\beta, e^o) - u(\beta_N, e^o) = 0$, the desired result follows. □

We use this lemma to prove Proposition 2 by contradiction. Suppose to the contrary that the optimal bonus scheme takes a non-hurdle form β_N and elicits effort e^* from the agent. Due to the enforcement condition, the scheme must obey $0 \leq \beta_N(\mathbf{x}) \leq \frac{\delta}{1-\delta} s(e^*)$ for all $\mathbf{x}$. Then it follows from Lemma 1 that there exists a hurdle scheme $\beta(\mathbf{x}) = \frac{\delta}{1-\delta} s(e^*) \mathbb{1}_{\{l(\mathbf{x}, e^*) \geq \kappa\}}(\mathbf{x})$ for which $u(\beta, e^*) = u(\beta_N, e^*)$ but $u_e(\beta, e^*) > u_e(\beta_N, e^*)$.

Observe that if the constructed hurdle scheme β satisfies

$$u(\beta, e) \leq u(\beta_N, e) \quad \text{for all } e \leq e^*, \tag{12}$$

then it implements higher effort than e^* than β_N, contradicting that β_N is the optimal contract. We now demonstrate that (12) is indeed true whenever the signal $\mathbf{X}$ satisfies the MLRTP.

Suppose that $u(\beta, e') > u(\beta_N, e')$ for some $e' < e^*$. Along with $u(\beta, e^*) = u(\beta_N, e^*)$ and $u_e(\beta, e^*) > u_e(\beta_N, e^*)$, it follows from the mean value theorem that there exists an $e^0 \in (e', e^*)$ satisfying $u(\beta, e^o) = u(\beta_N, e^o)$ and $u_e(\beta, e^o) < u_e(\beta_N, e^o)$. However by the MLRTP, the hurdle scheme β can be reformulated as

$$\beta(\mathbf{x}) = \frac{\delta}{1-\delta} s(e^*) \mathbb{1}_{\{l(\mathbf{x}, e^*) < \kappa\}}(\mathbf{x}) = \frac{\delta}{1-\delta} s(e^*) \mathbb{1}_{\{l(\mathbf{x}, e^o) < \kappa'\}}(\mathbf{x}),$$

i.e., a hurdle scheme for the likelihood ratio $l(\mathbf{x}, e^o)$ with the adjusted hurdle κ'. Then the same argument in the proof of Lemma 1 can be used to show $u_e(\beta, e^o) \geq u_e(\beta_N, e^o)$, thereby establishing (12) and optimality of the hurdle scheme. The proof is now complete.

A.3. Proof of Proposition 3

Suppose to the contrary that the optimal bonus b is strictly lower than $b^* = \frac{\delta}{1-\delta}s(e^*)$. Since the optimal effort $e^* \leq e^{FB}$ is assumed to be an interior solution, (IC_L) must hold at $e = e^*$ and thus we have

$$0 \;=\; b\int_{\kappa<l(\mathbf{x},e^*)} l(\mathbf{x}, e^*)f(\mathbf{x}, e^*)d\mathbf{x} - c'(e^*). \tag{13}$$

From this condition, it is straightforward to see that a higher bonus than b will enhance the agent's marginal incentive for effort at e^*. Moreover, in light of the stochastic dominance result in Proposition 1, the higher bonus will relax all downward IC constraints (IC_GD) because

$$c(e^*) - c(e) \;\leq\; b\Big[\underbrace{\Pr(l(\mathbf{X}, e^*) < \kappa \,|\, e) - \Pr(l(\mathbf{X}, e^*) < \kappa \,|\, e^*)}_{\geq 0 \;\text{ for all } e \leq e^*}\Big]$$

Accordingly, the higher bonus will induce higher effort than b, and thus the scheme with $b < b^*$ cannot be optimal.

To prove part (ii) of Proposition 3, suppose that the ratio $l(\mathbf{x}, e)$ is a decreasing function in e for all $\mathbf{x}$ but the optimal hurdle κ is strictly positive. Recall that the optimal effort e^* is the highest one satisfying (IC_GD) and the local IC constraint (13) must hold at e^*, with $b = b^*$.

Now consider a more lenient hurdle κ' in $(0, \kappa)$. Then it is immediate from (13) that κ' will increase the agent's marginal incentives. It remains to check if the scheme with hurdle κ' satisfies (IC_GD). To this end, it is enough to show that for all $e \leq e^*$

$$\Pr(l(\mathbf{X}, e^*) < \kappa \,|\, e) - \Pr(l(\mathbf{X}, e^*) < \kappa \,|\, e^*) \;\leq\; \Pr(l(\mathbf{X}, e^*) < \kappa' \,|\, e) - \Pr(l(\mathbf{X}, e^*) < \kappa' \,|\, e^*),$$

or to put it another way,

$$\Pr(l(\mathbf{X}, e^*) < \kappa \,|\, e) - \Pr(l(\mathbf{X}, e^*) < \kappa' \,|\, e) \;\text{ increases with } e \leq e^*. \tag{14}$$

We demonstrate that if $l_e(\mathbf{x}, e) \leq 0$ for all $\mathbf{x}$, then (14) is the case for all $0 < \kappa' < \kappa$. For this purpose, we take the derivative of the expression in (14) with respect to e to obtain

$$\int_{\kappa'<l(\mathbf{x},e^*)\leq\kappa} l(\mathbf{x}, e)f(\mathbf{x}, e)d\mathbf{x} \;>\; \kappa'\int_{\kappa'<l(\mathbf{x},e^*)\leq\kappa} f(\mathbf{x}, e)d\mathbf{x},$$

where the inequality results from $l(\mathbf{x}, e) \geq l(\mathbf{x}, e^*)$ for all $e \leq e^*$ and $l(\mathbf{x}, e^*) > \kappa'$. This ensures that the sign of the derivative is positive, and thus the scheme with κ' alleviates all downward IC constraints. This establishes that the optimal bonus scheme must involve a nonpositive hurdle. $\kappa \geq 0$ if $l_e \geq 0$ can be shown in an analogous way.

A.4. Proof of Proposition 4

For the sufficiency part, it is enough to show that if $\mathbf{X} \succ_L \mathbf{Y}$, then the optimal effort under signal $\mathbf{Y}$ is implementable by a contract based on signal $\mathbf{X}$.

Choose an agency problem $\omega = \langle v, c, \overline{\pi}, \overline{u}, \delta \rangle$ from the set Ω, and denote by $e_{\mathbf{Y}}$ the optimal effort given signal $\mathbf{Y}$ in this problem. In light of Proposition 2 and 3, we see that there exists a hurdle scheme

$$\beta(\mathbf{y}) = \begin{cases} 0 & \text{if } \ l(\mathbf{y}, e_{\mathbf{Y}}) < \kappa_{\mathbf{Y}} \\ b_{\mathbf{Y}} & \text{otherwise,} \end{cases}$$

with $b_{\mathbf{Y}} \equiv \frac{\delta}{1-\delta} s(e_{\mathbf{Y}})$, which implements $e_{\mathbf{Y}}$. Hence the given contract $(w, \beta, e_{\mathbf{Y}})$ must satisfy the global IC constraint: $u(\beta, e_{\mathbf{Y}}) \geq u(\beta, e)$ for all e, or equivalently

$$L_{\mathbf{Y}}(\kappa_{\mathbf{Y}}, e) - L_{\mathbf{Y}}(\kappa_{\mathbf{Y}}, e_{\mathbf{Y}}) \ \geq \ \frac{c(e_{\mathbf{Y}}) - c(e)}{b_{\mathbf{Y}}} \quad \text{for all } \ e \in [0, \bar{e}].$$

Then it follows from $\mathbf{X} \succ_L \mathbf{Y}$ that there exists a hurdle κ' for signal $\mathbf{X}$ such that

$$L_{\mathbf{X}}(\kappa', e) - L_{\mathbf{X}}(\kappa', e_{\mathbf{Y}}) \ \geq \ \frac{c(e_{\mathbf{Y}}) - c(e)}{b_{\mathbf{Y}}} \quad \text{for all } \ e \leq e_{\mathbf{Y}}.$$

The inequality implies that the optimal effort $e_{\mathbf{Y}}$ under $\mathbf{Y}$ is also implementable under signal $\mathbf{X}$ by a hurdle scheme β' for the likelihood ratio $l(\mathbf{x}, e_{\mathbf{Y}})$, which awards fixed bonus $b_{\mathbf{Y}}$ iff $l(\mathbf{x}, e_{\mathbf{Y}}) > \kappa'$. This proves that $\mathbf{X} \succ_L \mathbf{Y}$ is sufficient for $\mathbf{X}$ to be a more informative signal than $\mathbf{Y}$ within Ω.

To prove the converse, suppose to the contrary that $\mathbf{X}$ does not dominate $\mathbf{Y}$ in the likelihood ratio order. This means that there exists a κ and e^* such that for all $\kappa' \in \Re$,

$$L_{\mathbf{Y}}(\kappa, e_\kappa) - L_{\mathbf{Y}}(\kappa, e^*) \ > \ L_{\mathbf{X}}(\kappa', e_\kappa) - L_{\mathbf{X}}(\kappa', e^*) \quad \text{for some } \ e_\kappa < e^*. \tag{15}$$

Below we demonstrate that if the two signals are not ranked in the likelihood ratio order, then there exist a pair of contracting parties $\omega \in \Omega$ for which e^* is implementable under signal $\mathbf{Y}$, whereas no effort $e \geq e^*$ is implementable under signal $\mathbf{X}$. This contradicts with our assumption that $e_{\mathbf{X}}(\omega) \geq e_{\mathbf{Y}}(\omega)$ for all $\omega \in \Omega$.

The proof is by construction. To this end, denote by Γ the set of decreasing C^1-functions defined on the compact set $[0, e^*]$ such that $\gamma(e^*) = 0$ for all $\gamma \in \Gamma$. With κ and e^* that are specified in (15), note that $L_{\mathbf{Y}}(\kappa, e) - L_{\mathbf{Y}}(\kappa, e^*)$, regarding as a function of e and restricting its domain to $[0, e^*]$, is an element of Γ due to the stochastic dominance property of the MLRTP. Similarly, for every $\kappa' \in \Re$, we have $L_{\mathbf{X}}(\kappa', e) - L_{\mathbf{X}}(\kappa', e^*) \in \Gamma$. Observe that if (15) holds, we can select a function γ from Γ satisfying

$$L_{\mathbf{Y}}(\kappa, e) - L_{\mathbf{Y}}(\kappa, e^*) \ \geq \ \gamma(e) \ \text{ for all } \ e \leq e^* \quad \text{and} \tag{16}$$

$$L_{\mathbf{X}}(\kappa', e_\kappa) - L_{\mathbf{X}}(\kappa', e^*) \ < \ \gamma(e_\kappa) \ \text{ for } \ e_\kappa < e^*. \tag{17}$$

To construct the function γ on the remaining domain $[e^*, \bar{e}]$, we first fix $\gamma(e^*) = 0$ for its continuity and choose a continuous decreasing function satisfying

$$\gamma(e) \;\le\; \min\left\{L_{\mathbf{Y}}(\kappa', e) - L_{\mathbf{Y}}(\kappa', e^*),\; \min_{\kappa \in \Re}\; \left\{L_{\mathbf{X}}(\kappa, e) - L_{\mathbf{X}}(\kappa, e^*)\right\}\right\} \quad \forall\, e \in [e^*, \bar{e}], \tag{18}$$

with the inequality being strict except at $e = e^*$, where the min operator inside the curly bracket indicates the point minimization at each e.

We are now ready to construct the agent's cost function $c(e)$. We first assign one positive value to $c(e^*) \ge \frac{\delta}{1-\delta} s(e^*)\gamma(0)$ so that the cost function defined below takes a nonnegative value everywhere on $[0, \bar{e}]$:

$$c(e) \;=\; \begin{cases} c(e^*) - \frac{\delta}{1-\delta} s(e^*)\,\gamma(e) & \text{for } \; e \in [0, e^*] \\ c(e^*) - \frac{\delta}{1-\delta} s(e^{FB})\,\gamma(e) & \text{for } \; e \in [e^*, \bar{e}], \end{cases}$$

where the principal's objective v, an increasing and C^1 function, and the parties' reservation payoffs are chosen from the set Ω such that (i) the expected surplus $s(e) = v(e) - c(e) - \bar{\pi} - \bar{u}$ attains its maximum at $e^{FB} > e^*$ and is increasing on $(0, e^{FB})$, (ii) $s(e^*) > 0$, and (iii) the defined cost function is differentiable at $e = e^*$, that is,

$$\lim_{e \downarrow e^*} c'(e) \;=\; -\frac{\delta}{1-\delta} s(e^{FB}) \lim_{e \downarrow e^*} \gamma'(e) \;=\; -\frac{\delta}{1-\delta} s(e^*) \lim_{e \uparrow e^*} \gamma'(e) \;=\; \lim_{e \uparrow e^*} c'(e).$$

Then the constructed problem $\omega = \langle v, c, \bar{\pi}, \bar{u}, \delta \rangle$ is an element of Ω.

We now demonstrate that in the problem ω, e^* is implementable under signal $\mathbf{Y}$ but neither e^* nor any higher effort than e^* is implementable under signal $\mathbf{X}$. For $\mathbf{Y}$, consider the hurdle scheme β, $\beta(\mathbf{y}) = 0$ if $l(\mathbf{y}, e^*) < \kappa$ and $\beta(\mathbf{y}) = \frac{\delta}{1-\delta} s(e^*)$ otherwise, where κ and e^* are defined as in (15). Then it follows from (16) that the agent would find any downward deviation $e \le e^*$ nonprofitable. No upward deviation $e > e^*$ is profitable either, since (18) yields

$$u(\beta, e) - u(\beta, e^*) \;=\; \frac{\delta}{1-\delta} s(e^*) \Big[L_{\mathbf{Y}}(\kappa, e^*) - L_{\mathbf{Y}}(\kappa, e)\Big] + \frac{\delta}{1-\delta} s(e^{FB})\gamma(e) \;<\; 0.$$

Therefore, e^* is implementable with a feasible hurdle scheme under signal $\mathbf{Y}$.

In the same manner as above, it can be shown that e^* is not implementable with any hurdle scheme under $\mathbf{X}$. To see that any higher target effort than e^* is subject to deviations, suppose to the contrary that some effort $e_{\mathbf{X}} \in (e^*, e^{FB}]$ is implementable under $\mathbf{X}$. By Proposition 2 and 3, amongst possible payment schemes implementing $e_{\mathbf{X}}$, we can restrict attention to a hurdle scheme, $\beta(\mathbf{x}) = 0$ for $l(\mathbf{x}, e_{\mathbf{X}}) < \kappa_{\mathbf{X}}$, or equivalently by the MLRTP, $\beta(\mathbf{x}) = 0$ for $l(\mathbf{x}, e^*) < \kappa^*$, and $\beta(\mathbf{x}) = \frac{\delta}{1-\delta} s(e_{\mathbf{X}})$ for the other case. But given any schemes taking this form, the agent would deviate by choosing $e^* < e_{\mathbf{X}}$ because

$$u(\beta, e^*) - u(\beta, e_{\mathbf{X}}) \;=\; \frac{\delta}{1-\delta} s(e_{\mathbf{X}}) \Big[L_{\mathbf{X}}(\kappa^*, e_{\mathbf{X}}) - L_{\mathbf{X}}(\kappa^*, e^*)\Big] - \frac{\delta}{1-\delta} s(e^{FB})\gamma(e_{\mathbf{X}}) \;>\; 0,$$

where the strict inequality follows from $s(e_{\mathbf{X}}) \leq s(e^{FB})$ and the inequality assumed in (18), which holds with a strict inequality at $e = e_{\mathbf{X}} > e^*$. Therefore, no feasible schemes under signal $\mathbf{X}$ can implement effort $e \geq e^*$. This contradiction establishes that $\mathbf{X} \succ_L \mathbf{Y}$ is necessary for X to be more informative than $\mathbf{Y}$ within class Ω. The proof is now complete.

A.5. Proof of Proposition 5

Given a pair of multivariate signals $\mathbf{X}$ and $\mathbf{Y}$, suppose that the likelihood ratio of $\mathbf{X}$ is more precise than the ratio of $\mathbf{Y}$ in the sense of (P). That is, for each constant κ, there exists an increasing function $T_\kappa : [0, \bar{e}] \to \Re$ such that

$$L_{\mathbf{X}}(T_\kappa(e), e) \;=\; L_{\mathbf{Y}}(\kappa, e) \qquad \forall\, e \in [0, \bar{e}]. \tag{19}$$

Given the constant κ and effort e^*, let $\kappa' \equiv T_\kappa(e^*)$. Then we have

$$\begin{aligned} L_{\mathbf{Y}}(\kappa, e) - L_{\mathbf{Y}}(\kappa, e^*) \;&=\; L_{\mathbf{X}}(T_\kappa(e), e) - L_{\mathbf{X}}(\kappa', e^*) \\ &\leq\; L_{\mathbf{X}}(\kappa', e) - L_{\mathbf{X}}(\kappa', e^*), \end{aligned}$$

where the inequality results from monotonicity of the T-transformation in (19): $T_\kappa(e) \leq T_\kappa(e^*) = \kappa'$. As κ was arbitrarily chosen, the inequality established above implies $\mathbf{X} \succ_L \mathbf{Y}$. This proves part (i) of Proposition 5.

To prove part (ii), assume that two unidimensional signals, $X \sim F$ and $Y \sim G$, satisfy the strict MLRP. So far as we prove this part, in order to avoid confusions, we use subscript X and Y to denote the likelihood ratios and CDFs. By the strict MLRP, both ratios $l_X(\cdot, e^*)$ and $l_Y(\cdot, e^*)$ are strictly increasing in the first argument. We demonstrate that in this case, $l_X(\cdot, e^*)$ is more precise than $l_Y(\cdot, e^*)$ if and only if X is more precise than Y.[30]

Suppose X is more precise than Y. Define by λ the inverse function of $l_Y(\cdot, e^*)$ with respect to the first argument, and let $\kappa = l_Y(y, e^*)$ for some y in the support of signal Y, or equivalently $y = \lambda(\kappa)$. Then it follows that

$$L_Y(\kappa, e) \;=\; \Pr(l_Y(Y, e^*) \leq \kappa |\, e) \;=\; G(\lambda(\kappa), e) \;=\; F\left(T_{\lambda(\kappa)}(e), e\right), \tag{20}$$

where the last equality is due to $X \succ_{\mathrm{P}} Y$. Let $H_\kappa(e) \equiv l_X(T_{\lambda(\kappa)}(e), e^*)$. Since the inner function $T_{\lambda(\kappa)}$ increases with e and the outer function l_X increases with the first argument, their composition $H_\kappa(e)$ is an increasing function of e. Then by definition of L_X and $H_\kappa(e)$, we have

$$L_X\left(H_\kappa(e), e\right) \;=\; \Pr\left(l_X(X, e^*) \leq l_X\left(T_{\lambda(\kappa)}(e), e^*\right) \,\middle|\, e\right) \;=\; F\left(T_{\lambda(\kappa)}(e), e\right).$$

Putting together with (20), we obtain $L_Y(\kappa, e) = L_X(H_\kappa(e), e)$ which establishes $l_X \succ_{\mathrm{P}} l_Y$.

To prove the converse, suppose that $l_X \succ_{\mathrm{P}} l_Y$, i.e., for every κ, there exists an increasing

[30] A similar argument can be used to establish a more general result: X is more precise than Y if and only if $m(X)$ is more precise than $n(Y)$ with $m(\cdot)$ and $n(\cdot)$ strictly increasing.

function $H_\kappa(e)$ satisfying $L_X(H_\kappa(e), e) = L_Y(\kappa, e)$ for all e. Select a sample y from the support of Y, and put $\kappa(y) = l_Y(y, e^*)$. With $\kappa(y)$ and H_κ, we can write the CDF of Y as

$$G(y,e) \;=\; L_Y\big(l_Y(y,e^*),e\big) \;=\; L_Y\big(\kappa(y),e\big) \;=\; L_X\left(H_{\kappa(y)}(e),e\right). \tag{21}$$

Define by μ the inverse of $l_X(\cdot, e^*)$ with respect to the first argument. With μ we can write the distribution of l_X as

$$L_X(\kappa',e) \;=\; \Pr\left(l_X(X,e^*) \le \kappa' \,\middle|\, e\right) \;=\; F\left(\mu(\kappa'),e\right). \tag{22}$$

Putting (21) and (22) together leads us to

$$G(y,e) \;=\; F\left(\mu\left(H_{\kappa(y)}(e)\right),e\right) \quad \forall\, e.$$

Since both μ and H_κ are increasing in their arguments, the composite function $\mu(H_{\kappa(y)}(e))$ must increase with e. This establishes the existence of an increasing transformation associating F and G, and therefore $X \succ_P Y$. The proof is now complete. □

B. The regularity condition

In this section, we derive the critical ratio introduced by Poblete and Spulber (2012) in the contracting environment with a verifiable unidimensional signal and compare their regularity condition with the condition MLRTP.

Assume that the outcome $X \sim F(\cdot, e)$ with support $[\underline{x}, \overline{x}]$ indicates the principal's objective, and denote by $s(x)$ the sharing rule between the two risk-neutral parties. After carrying out a contract, the principal obtains a payoff of $x - s(x)$ and the agent obtains $s(x) - c(e)$. The critical ratio $\rho(x, e)$ is defined as the ratio of the net expected benefit to cost from increasing the slope of $s(x)$ by Δ over $[x, x + dx]$. The increment of $s'(x)$ increases the principal's expected benefit by $-\Delta F_e(x, e)dx$ through the agent's marginal incentive, and at the same time aggravates her expected cost by $\Delta(1 - F(x, e))dx$.[31] Therefore, the critical ratio reduces into

$$\rho(x,e) \;=\; \frac{-\Delta F_e(x,e)dx}{\Delta(1-F(x,e))dx} \;=\; -\frac{F_e(x,e)}{1-F(x,e)}.$$

The critical ratio is regular if $\rho(x', e) \ge \rho(x, e)$ for some x and x' and for some e implies $\rho(x', e') \ge \rho(x, e')$ for all e'. Observe that the defined regularity imposes the same condition on the critical ratio as what MLRTP imposes on the likelihood ratio.

[31] Integrating by parts, the agent's expected payoff from $s(x)$ can be written as

$$\int_{\underline{x}}^{\overline{x}} s(x)\, f(x,e)dx \;=\; \int_{\underline{x}}^{\overline{x}} s'(x)\Big(1-F(x,e)\Big)dx.$$

Therefore, the increment of $s'(x)$ by Δ over $[x, x + dx]$ would strengthen the agent's marginal incentives by $-\Delta F_e(x, e)dx$ and total incentives by $\Delta(1 - F(x, e))dx$, respectively.

To examine the relation between the two conditions, recall that if the density function $f(x,e)$ is log-supermodular, then the hazard rate $f(x,e)/[1-F(x,e)]$ is decreasing in e for all x. This in turn is equivalent to the increasing critical ratio property for all e. As a result, if $l(x,e)$ is increasing in x for all e, then $\rho(x,e)$ is increasing in x for all e as well. The reverse is not true, so MLRP is not always necessary for $\rho(x,e)$ to be increasing. See Poblete and Spulber (2012) for a counterexample. Hence the MLRTP is seemingly more restrictive than the regularity condition of $\rho(x,e)$. However, the following example shows that there exists a signal which satisfies the MLRTP but does not have regular critical ratios, suggesting that there is no inclusive relationship between them.

EXAMPLE 3. *Consider a signal $X \sim N(0,\sigma(e)^2)$, where the agent's effort does not affect the mean of X but affects its variance. This signal has likelihood ratio*

$$l(x,e) \;=\; \frac{\partial}{\partial e}\ln f(x,e) \;=\; \left[-1+\left(\frac{x}{\sigma(e)}\right)^2\right]\frac{\sigma'(e)}{\sigma(e)}.$$

Suppose $\sigma'(e)>0$ for all e, implying that as the agent exerts higher effort, the output distribution is more diffuse. Then for every admissible e, the ratio $l(x,e)$ is increasing in x for all $x \geq 0$ but decreasing for all $x<0$, and satisfies the MLRTP.

On the other hand, the signal X has the critical ratio as

$$\rho(x,e) \;=\; \frac{z\phi(z)}{1-\Phi(z)}\cdot\frac{\sigma'(e)}{\sigma(e)}, \qquad z=\frac{x}{\sigma(e)},$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ indicate the c.d.f and p.d.f of the standard normal distribution, respectively. Then

$$\frac{\partial}{\partial x}\rho(x,e) \;=\; \rho_x(x,e) \;=\; \frac{d}{dz}\left(\frac{z\phi(z)}{1-\Phi(z)}\right)\frac{\sigma'(e)}{\sigma(e)^2},$$

where

$$\begin{aligned}\frac{d}{dz}\left(\frac{z\phi(z)}{1-\Phi(z)}\right) &= \frac{1}{(1-\Phi(z))^2}\Big[(z\phi'(z)+\phi(z))(1-\Phi(z))-z\phi(z)(-\phi(z))\Big]\\ &= \frac{\phi(z)}{(1-\Phi(z))^2}\Big[(-z^2+1)(1-\Phi(z))+z\phi(z)\Big], \qquad \text{since } \phi'(z)=-z\phi(z),\\ &\equiv \Psi(z).\end{aligned}$$

Note that the derivative $\Psi(z)$ takes a strictly positive value at $z=0$ but a negative value at $z=-1$. Hence there exist $z_1<z_2<0$ such that $\Psi(z_1)<0$ and $\Psi(z_2)>0$.

For given e_1, let $x_1=z_1\sigma(e_1)$ and let $e_2>e_1$ denote the level of effort at which $\sigma(e_2)z_2=x_1$. Then we have

$$\rho_x(x_1,e_1) \;=\; \Psi(z_1)\cdot\frac{\sigma'(e_1)}{\sigma(e_1)^2} < 0 \quad \text{and} \quad \rho_x(x_1,e_2) \;=\; \Psi(z_2)\cdot\frac{\sigma'(e_2)}{\sigma(e_2)^2} \;>\; 0.$$

That is, the critical ratio is decreasing in x for effort e_1 but increasing for e_2. Consequently, $\rho(x,e)$ is not *regular.*

References

BANKS, D. (1963). An integral inequality. *Proceedings of the American Mathematical Society*, **14** (5), 823–828.

BLACKWELL, D. (1951). *Comparison of Experiments*. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press.

— (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, **24** (2), 265–272.

BULL, C. (1987). The existence of self-enforcing implicit contracts. *The Quarterly Journal of Economics*, **102** (1), 147–59.

CHI, C. K. and CHOI, K. J. (2018). *The Value of Information in Agency Models*. Tech. rep., Norwegian School of Economics (NHH).

CONLON, J. R. (2009). Two New Conditions Supporting the First-Order Approach to Multisignal Principal-Agent Problems. *Econometrica*, **77** (1), 249–278.

DEWATRIPONT, M., JEWITT, I. and TIROLE, J. (1999). The economics of career concerns, part i: Comparing information structures. *Review of Economic Studies*, **66** (1), 183–98.

GILLAN, S. L., HARTZELL, J. C. and PARRINO, R. (2009). Explicit versus Implicit Contracts: Evidence from CEO Employment Agreements. *Journal of Finance*, **64** (4), 1629–1655.

GJESDAL, F. (1982). Information and Incentives: The Agency Information Problem. *Review of Economic Studies*, **49** (3), 373–90.

GROSSMAN, S. J. and HART, O. D. (1983). An analysis of the principal-agent problem. *Econometrica*, **51** (1), 7–45.

HOLMSTRÖM, B. (1979). Moral hazard and observability. *Bell Journal of Economics*, **10** (1), 74–91.

HWANG, S. (2016). Relational contracts and the first-order approach. *Journal of Mathematical Economics*, **63** (C), 126–130.

INNES, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory*, **52** (1), 45–67.

JEWITT, I. (2007). *Information Order in Decision and Agency Problems*. Tech. rep., Working Paper.

JUNG, J. Y. and KIM, S. K. (2015). Information space conditions for the first-order approach in agency problems. *Journal of Economic Theory*, **160** (C), 243–279.

KE, R. and RYAN, C. T. (2018). A general solution method for moral hazard problems. *Theoretical Economics*, **13** (3).

KIM, S. K. (1995). Efficiency of an information system in an agency model. *Econometrica*, **63** (1), 89–102.

KIRKEGAARD, R. (2017a). A unifying approach to incentive compatibility in moral hazard problems. *Theoretical Economics*, **12** (1), 25–51.

— (2017b). Moral hazard and the spanning condition without the first-order approach. *Games and Economic Behavior*, **102** (C), 373–387.

KLEIN, B. and LEFFLER, K. B. (1981). The Role of Market Forces in Assuring Contractual Performance. *Journal of Political Economy*, **89** (4), 615–641.

KVALØY, O. and OLSEN, T. E. (2014). *Teams and Tournaments in Relational Contracts*. CESifo Working Paper Series 4783, CESifo Group Munich.

— and OLSEN, T. E. (2017). *Relational Contracts and Multitasking*. Tech. rep., Mimeo.

LAZEAR, E. and ROSEN, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, **89** (5), 841–64.

LEHMANN, E. L. (1988). Comparing location experiments. *The Annals of Statistics*, **16** (2), 521–533.

LEVIN, J. (2003). Relational incentive contracts. *American Economic Review*, **93** (3), 835–857.

MACAULAY, S. (1963). Non-contractual relations in business: A preliminary study. *American Sociological Review*, **28** (1), 55–67.

MACLEOD, B. W. (2003). Optimal Contracting with Subjective Evaluation. *American Economic Review*, **93** (1), 216–240.

MACLEOD, W. B. and MALCOMSON, J. M. (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica*, **57** (2), 447–80.

MALCOMSON, J. M. (2012). Relational Incentive Contracts. In R. Gibbons and J. Roberts (eds.), *The Handbook of Organizational Economics*, Introductory Chapters, Princeton University Press.

MILGROM, P. R. (1981). Good News and Bad News: Representation Theorems and Applications. *Bell Journal of Economics*, **12** (2), 380–391.

— (1988). Employment Contracts, Influence Activities, and Efficient Organization Design. *Journal of Political Economy*, **96** (1), 42–60.

MILLER, D., OLSEN, T. E. and WATSON, J. (2018). *Relational Contracting, Negotiation, and External Enforcement*. Discussion Papers 2018/8, Norwegian School of Economics, Department of Business and Management Science.

MIRRLEES, J. (1979). The implications of moral hazard for optimal insurance. In *Seminar given at Conference held in honor of Karl Borch, Bergen, Norway*.

MURPHY, K. J. and OYER, P. (2003). *Discretion in Executive Incentive Contracts: Theory and Evidence*. Tech. rep., Stanford University.

PERSICO, N. (2000). Information acquisition in auctions. *Econometrica*, **68** (1), 135–148.

POBLETE, J. and SPULBER, D. (2012). The form of incentive contracts: agency with moral hazard, risk neutrality, and limited liability. *RAND Journal of Economics*, **43** (2), 215–234.

PRENDERGAST, C. and TOPEL, R. H. (1996). Favoritism in Organizations. *Journal of Political Economy*, **104** (5), 958–978.