A Proxy-Contract Based Approach to the First-Order Approach in Agency Models*

Jin Yong Jung† Son Ku Kim‡ Seung Joo Lee§

November 14, 2023

Abstract

To justify the first-order approach in principal-agent problems, the previous literature has focused on making the agent's expected monetary utility obtained from this approach *concave* in the agent's effort. However, relying on such concavity is overly sufficient. We propose new sets of conditions based on a novel double-crossing property between a 'proxy' contract and the optimal contract derived from the first-order approach, extending the applicability of the first-order approach significantly. Due to the flexibility in choosing a proxy contract, our new approach can be applied to a wider range of principal-agent problems, in which the previous literature does not justify the use of the first-order approach.

*We appreciate John Conlon for his extensive comments. This paper was previously circulated with the title 'Justifying the First-Order Approach in Agency Frameworks with the Agent's Possibly Non-Concave Value Function'.

†Kangnam University (jyjung@kangnam.ac.kr)

‡Seoul National University (sonkukim@snu.ac.kr)

§Saïd Business School, Oxford University (seung.lee@sbs.ox.ac.uk)

1 Introduction

Replacing the agent's original 'argmax' incentive compatibility constraint with its first-order condition with respect to his effort, which is called the first-order approach, has been typically adopted in solving the principal-agent problems. However, this approach is not always valid even in the standard setting.[1] Therefore, identifying conditions under which relying on the first-order approach is valid has been one of the major issues in the literature on principal-agent problems, and several sets of sufficient conditions for justifying this approach have been found in various agency frameworks.[2]

A common approach employed by the previous literature is to find sufficient conditions under which the agent's 'expected' monetary utility obtained from this approach becomes concave in his effort, ensuring that the agent's original incentive compatibility is satisfied at the designated effort level. As the agent's expected monetary utility obtained from the first-order approach depends on both the characteristics of distributions of relevant signals (i.e., technology) and the agent's preference, the previous literature can be broadly classified into two branches: while the literature's first branch focuses on sufficient conditions on distributions only (see e.g., Mirrlees (1975), Rogerson (1985), Sinclair-Desgagné (1994), Conlon (2009), and Jung and Kim (2015)), the second sub-literature imposes conditions both on signals' distributions and the agent's utility function (see e.g., Jewitt (1988) and Jung and Kim (2015)). For example, the requirements only imposed on distributions can be too strict to be satisfied by large classes of distributions. Overcoming this issue, the literature's second branch imposes weaker conditions on distributions but puts additional restrictions on the agent's preference instead, so that broader classes of distributions can be used in the principal-agent problems for the purpose of using the first-order approach. Yet still in some cases, those conditions can be quite strict so that there are many meaningful principal-agent settings in which using the first-order approach cannot be justified by the literature's existing sets of conditions: for example, in cases where the optimal contract is to be bounded below due to the agent's limited liability, so is his indirect utility given that optimal contract, and therefore it cannot be concave in the likelihood ratio, satisfying neither the Jewitt (1988)'s conditions nor the Jung and Kim (2015)'s conditions.

In this paper, we present a completely different approach to this long-standing problem.

[1]For the standard principal-agent framework, see Spence and Zeckhauser (1971), Ross (1973), Mirrlees (1975), Harris and Raviv (1979), Holmstrom (1979), Shavell (1979), and Grossman and Hart (1983), among others.

[2]For a detailed review of the literature on the first-order approach, see e.g., Jung and Kim (2015).

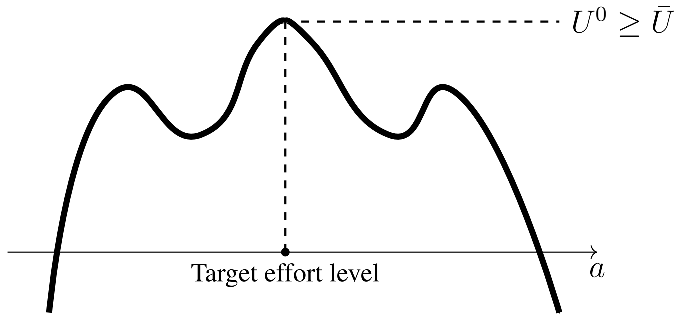


Figure 1: Possibly Non-Concave Indirect Monetary Utility of the Agent

Actually, our approach is not at all based on making the agent's expected monetary utility under the optimal contract derived from the first-order approach 'concave' in his effort. For example, as shown in Figure 1, the first-order approach can be justified as long as the agent's expected utility from this approach has a maximum value at a target effort level the principal intends to induce from the agent (i.e., the original argmax incentive constraint is satisfied). Given any target effort level, our approach can be broadly summarized as follows: (i) we come up with a 'proxy' contract under which the agent takes the same target effort (i.e., given the proxy contract, the agent's expected monetary utility subtracted by his cost of action is maximized at the given target effort level); (ii) We show that the agent's expected monetary utility under the optimal contract based on the first-order approach (as a function of the agent's effort choice) is always below that under the proxy contract, except when the agent takes the target effort level. Only when the agent takes the target effort, the expected monetary utility levels under the two contracts (i.e., the proxy contract and the optimal contract based on the first-order approach) coincide; (iii) Therefore, the agent's expected monetary utility under the optimal contract based on the first-order approach gets maximized at the target effort, satisfying the agent's original incentive compatibility constraint.

Under our approach, the agent's expected monetary utilities under both contracts need not be concave in the agent's effort: conditions we impose are that (i) the agent's expected monetary utility under the contrived proxy contract as a function of his action is single-peaked at the target effort level; and (ii) his expected monetary utility under the first-order approach is lower than under the proxy contract, except at the target effort level at which two become the same. And then we provide novel ways to contrive those proxy contracts, based on so-called 'double-crossing' properties between the agent's indirect utility functions under those two contracts (i.e., the optimal contract based on the first-order approach and the proxy one).

Our approach is fundamentally flexible, as we can pick various types of 'proxy' contracts with which we compare the optimal contract based on the first-order approach. This flexibility allows us to handle many meaningful cases that the previous literature does not clearly fit in: first, our approach can justify the first-order approach with distributions whose likelihood ratio is unbounded (e.g., the normal distributions) so that the agent's limited liability constraint is usually imposed for the solution's existence.[3] Second, our approach imposes far weaker conditions on the agent's preference than e.g., Jewitt (1988) and Jung and Kim (2015), thereby allowing that the agent's indirect utility as a function of the signals' likelihood ratio to be even convex. Based on this approach, our new sets of conditions contain a common interesting statistical condition on the signals' density functions. We illustrate that this statistical condition not only has a very useful implication but also is quite general in that it can be satisfied by the wide range of familiar density functions including the normal distributions and various other exponential families, whereas the literature's existing conditions cannot. We provide four alternative sets of conditions: three that can be used for cases in which the agent's limited liability constraint is not binding at the optimum, and one that can be used for cases in which the agent's limited liability constraint is binding for some signal values at the optimum, respectively. Under our proxy-contract based approach, the first-order approach can be validly adopted in many useful principal-agent settings (including the above normal distribution case) in which it has not been able to be justified by the literature's existing sets of conditions.

Related Literature The first-order approach problem has been known for a long time and various attempts to justify its use in many principal-agent settings have been made. The first set of conditions was proposed by Mirrlees (1975) and Rogerson (1985) in the one-signal case. Those conditions are the well-known MLRP (i.e., monotone likelihood ratio property) and the CDFC (i.e., convexity of the distribution function condition) for the distribution function of the signal. Later, other conditions were proposed to generalize the one-signal CDFC to multi-signal cases. They include the generalized CDFC (i.e., GCDFC) by Sinclair-Desgagné (1994), the CISP (i.e., concave increasing set probability condition)

[3]This is the well-known Mirrlees' unpleasant theorem (see e.g., Mirrlees (1975)). Also, see Jewitt et al. (2008) for the solution's existence and uniqueness in the presence of the agent's limited liability constraint.

by Conlon (2009), and the CDFCL (i.e., convexity of the distribution function condition for the likelihood ratio) by Jung and Kim (2015) among others.[4] All these conditions contain the property of the CDFC.

However, the CDFC and its various extensions have a serious limitation in that they are hardly satisfied by most familiar density functions. For instance, consider a one-signal principal-agent problem in which the signal for the agent's hidden effort is generated by a simple functional form such that $x = a + \theta$, where $a \in [0, \infty)$ is the agent's effort level, and $x \in \mathbb{R}$ and $\theta \in \mathbb{R}$ are the realized values of the signal variable, $\tilde{x}$, and the uncertainty variable, $\tilde{\theta} \sim N(0, \sigma^2)$, respectively.[5] It is well known that the density function of the signal conditional on the agent's effort in this case satisfies none of the above CDF-type conditions.

To overcome this drawback, Jewitt (1988) proposed another set of conditions in the one-signal case which does not contain any CDF-type condition, and is thus more applicable in general.[6] Recently, this set of conditions in Jewitt (1988) was also extended to the multi-signal case by Jung and Kim (2015).[7] These sets of conditions, however, have a different kind of limitation, and thus cannot be used for many familiar cases including the above normal density example. Although they do not contain any of the troublesome CDF-type conditions, they contain another condition that the agent's indirect utility given the optimal contract must be concave in the signals' likelihood ratio, which cannot be satisfied in many cases, including those that impose the agent's limited liability constraint.

To justify the first-order approach, the existing sets of conditions containing the CDF-type conditions (e.g., Rogerson (1985), Sinclair-Desgagné (1994), Conlon (2009), and Jung and Kim (2015)) were derived by putting all the requirements only on the signals' density function. Therefore, those CDF-type conditions are, in general, too restrictive to be satisfied by most familiar density functions. On the other hand, Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015), based on placing an additional requirement on the agent's utility function that the agent's indirect utility be concave in the likelihood ratio of the signals, were able to impose weaker requirements on the density function of the signals than those CDF-type conditions. However, there are still many cases in which such an ad-

[4] Jewitt (1988) also proposed two sets of conditions for the multi-signal case assuming that the multiple signals are independently distributed.

[5] We use letters with a tilde (e.g., $\tilde{x}$) to denote random variables and letters without it (e.g., x) to denote specific realized values of those random variables.

[6] See Theorem 1 in Jewitt (1988).

[7] See Proposition 7 in Jung and Kim (2015).

ditional requirement cannot be met. For many density functions of the signals (including the normal density function in the above example), their likelihood ratios are unbounded below. In this case, in order to guarantee the existence of the optimal contract, one needs to include the agent's limited liability constraint which requires that the agent's wage not be lower than a certain level under any circumstances. However, when the optimal contract is to be bounded below due to the agent's limited liability, so is the agent's indirect utility given that optimal contract, and thereby it cannot be globally concave in the likelihood ratio, satisfying neither the Jewitt (1988) conditions nor the Jung and Kim (2015) conditions. In this paper, we try to extend the applicability of the first-order approach in broader situations including those subtle cases, by relying on the totally different approach based on the comparison with a proxy contract and the double-crossing properties between the agent's indirect utility functions.

The remainder of the paper is organized as follows. In Section 2, we formulate the basic principal-agent framework, and briefly explain the issue of using the first-order approach. In Section 3, we present our main set of conditions that is based on our proxy-contract approach. Then, in Sections 4, we propose three alternative sets of conditions that are easier to verify than our main set of conditions in both the case where it is not binding at the optimum (i.e., Section 4.1) and the case where the agent's limited liability constraint is binding for some values of the signals at the optimum (i.e., Section 4.2), respectively. Concluding remarks are given in Section 5, and all formal proofs are relegated to the Appendix A. Appendix B contains the statistical implications of our sets of conditions, and compares them with the existing conditions, including those based on the concept of TP_3.

2 The Basic Model

We consider a one-period standard principal-agent model in which an agent works for a principal by inputting his effort $a \in [0, \overline{a}]$. The principal cannot observe the agent's effort choice directly but can observe some other variables $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ that are imperfectly correlated with the agent's hidden effort, where $\tilde{x}_i$ is a one-dimensional random variable. By taking the Mirrlees (1975) formulation, we denote $f(\mathbf{x}|a)$ as the joint density function of $\tilde{\mathbf{x}}$ conditional on the agent's effort, a. It is defined from the cumulative distribution function of $\tilde{\mathbf{x}}$ given a, i.e., $F(\mathbf{x}|a) \equiv Pr[\tilde{\mathbf{x}} \leq \mathbf{x}|a]$, where $\mathbf{x} \in \mathbb{R}^n$ is the realized value of signal vector $\tilde{\mathbf{x}}$. We assume that the support of $f(\mathbf{x}|a)$ is independent of a, and both $F(\mathbf{x}|a)$ and $f(\mathbf{x}|a)$ are continuously differentiable at least twice with respect to a, i.e.,

$F, f \in \mathbb{C}^2$.

When signal $\mathbf{x}$ is realized, the principal obtains $\pi(\mathbf{x})$ as the total value of the relationship with the agent, and she pays to the agent his wage s which depends on $\mathbf{x}$, i.e., $s = s(\mathbf{x})$. The principal is risk-neutral, whereas the agent is risk-averse. It is assumed that the agent's utility function takes an additively separable form such as

$$u(s, a) = u(s) - a, \;\; u' > 0, \;\; u'' < 0,$$

where $u(s)$ denotes the agent's utility from monetary payoff s.[8] Thus, the agent's expected utility when he takes an effort a under $s(\mathbf{x})$ is given by

$$U(s(\cdot), a) \equiv \int u(s(\mathbf{x})) f(\mathbf{x}|a) d\mathbf{x} - a. \tag{1}$$

We also assume that the agent can get $\overline{U}$ at maximum by working for other principals, thereby $\overline{U}$ is his reservation utility level. Furthermore, there is a limited liability constraint on the agent's side which requires that the agent's wage not be lower than $\underline{s}$ under any circumstances, i.e., $s(\mathbf{x}) \geq \underline{s}$, $\forall \mathbf{x}$, where $\underline{s}$ indicates the agent's subsistence wage level.[9] In the paper, it is assumed that $\overline{U} \geq u(\underline{s})$.

So a specific principal-agent problem can be represented by its characteristic variables $\{\pi(\mathbf{x}),u(s),f(\mathbf{x}|a),\overline{U},\underline{s}\}$, and the principal's optimization program with those variables is given by

$$\begin{aligned} \max_{a, s(\mathbf{x}) \geq \underline{s}} \qquad & \int [\pi(\mathbf{x}) - s(\mathbf{x})] f(\mathbf{x}|a) d\mathbf{x} \\ s.t. \quad & (i)\; U(s(\cdot), a) \geq \overline{U}, \\ & (ii)\; a \in \arg\max_{a'} \; U(s(\cdot), a'). \end{aligned}$$

In the above, the constraints are the typical participation and incentive compatibility constraints, respectively. This optimization program indicates that the principal has to decide both the agent's wage scheme, $s(\mathbf{x}) \geq \underline{s}$, $\forall \mathbf{x}$, and the target effort level, a, simultane-

[8]We use the prime and the double prime of a function to denote the first and the second derivatives of that function, respectively.

[9]The agent's limited liability constraint is especially needed to guarantee the existence of the optimal contract in some cases. Note that the case in which there is no limited liability constraint on the agent's side is a special case where $\underline{s} = -\infty$. For the existence issue, see Mirrlees (1975). Also, see Jewitt et al. (2008) for the solution's existence and uniqueness in the presence of the agent's limited liability constraint.

ously to maximize her expected payoff under the constraints that the self-interested agent actually chooses the target effort level when $s(\mathbf{x})$ is offered and that his expected utility in this case is not lower than $\overline{U}$.

However, the above program is generally not tractable in itself because the incentive constraint is composed of infinitely many inequality constraints. Thus, it has been typically solved by replacing the original incentive constraint with the *relaxed* constraint that the agent's expected utility given $s(\mathbf{x})$ is stationary at that effort level, a, i.e.,

$$\frac{\partial U(s(\cdot),a)}{\partial a} \equiv U_a(s(\cdot),a) = 0, \tag{2}$$

which is known as the first-order approach. Therefore, the principal's optimization program based on the first-order approach can be written as

$$\begin{aligned}\max_{a,s(\mathbf{x})\geq \underline{s}} \quad & \int [\pi(\mathbf{x}) - s(\mathbf{x})] f(\mathbf{x}|a) d\mathbf{x} \\ s.t. \quad & (i)\; U(s(\cdot),a) \geq \overline{U}, \\ & (ii)\; U_a(s(\cdot),a) = \int u(s(\mathbf{x})) f_a(\mathbf{x}|a) d\mathbf{x} - 1 = 0,\end{aligned}$$

where $f_a(\mathbf{x}|a) \equiv \frac{\partial f(\mathbf{x}|a)}{\partial a}$ is a partial derivative of the density function $f(\mathbf{x}|a)$ with respect to a.

Let $(s^o(\mathbf{x}), a^o > 0)$ solve the above optimization program.[10] By solving the Euler equation of the above program, one can derive that the optimal incentive contract, $s^o(\mathbf{x})$, should satisfy

$$\frac{1}{u'(s^o(\mathbf{x}))} = \begin{cases} \lambda + \mu \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}, & \text{if } s^o(\mathbf{x}) > \underline{s}, \\ \frac{1}{u'(\underline{s})}, & \text{otherwise}, \end{cases} \tag{3}$$

where λ and μ are the Lagrange multipliers of the participation and *relaxed* incentive compatibility constraints, respectively.

It is well known that the optimal contract, $s^o(\mathbf{x})$ in (3), does not always solve the original optimization program. This is because $s^o(\mathbf{x})$ in (3) is actually the optimal solution obtained by replacing the original "argmax" incentive constraint with the *relaxed* one, and the principal's opportunity set for $s(\mathbf{x})$ satisfying the relaxed incentive constraint is larger than her true opportunity set for $s(\mathbf{x})$ satisfying the original argmax incentive constraint,

[10] The existence of an optimal solution $(s^o(\mathbf{x}), a^o)$ is assumed. We also assume $a^o > 0$ to rule out a trivial case.

and thereby $s^o(\mathbf{x})$ sometimes may not be in her true opportunity set for $s(\mathbf{x})$. Thus, to guarantee that $s^o(\mathbf{x})$ in (3) actually solves the original program, it must be ensured that $s^o(\mathbf{x})$ satisfies the original argmax incentive constraint, that is,

$$U(s^o(\cdot), a) \leq U(s^o(\cdot), a^o), \quad \text{for all } \ a. \tag{4}$$

To ensure (4), all the existing results in the literature for justifying the first-order approach were derived to make the agent's expected utility as a function of a under $s^o(\mathbf{x})$, $U(s^o(\cdot), a)$, concave in a.[11] Obviously, given the concavity of $U(s^o(\cdot), a)$ in a, $U_a(s^o(\cdot), a^o) = 0$ guarantees (4), and thus using the first-order approach is valid. However, the concavity of $U(s^o(\cdot), a)$ in a is sufficient but not necessary for ensuring (4). As already drawn in Figure 1, even if $U(s^o(\cdot), a)$ is not concave, the first-order approach can still be justified. The main purpose of this paper is to find a new set of conditions which ensures (4) without relying on the concavity of $U(s^o(\cdot), a)$, and thus is more general than the existing sets of conditions.

3 Analysis

We start with proving that μ in (3) is positive. The basic proof for $\mu > 0$ was already given by Mirrlees (1975) and Holmstrom (1979).[12] But their proofs were given with the assumption that using the first-order approach is valid. Thus, it is obvious that those proofs cannot be used for finding conditions justifying the first-order approach itself. On the other hand, Jewitt (1988) provided another proof for $\mu > 0$ without such an assumption.[13] But, his approach also has a limitation in that it is valid only for the case in which the agent's limited liability constraint is not binding at the optimum. However, as will be shown later, there are actually many cases in which the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum, and one of our main goals is to provide the conditions under which using the first-order approach can be justified even in those cases. Therefore, we provide a different proof for $\mu > 0$ based neither on the assumption that using the first-order approach is valid nor on the assumption that the agent's limited liability constraint is not binding at the optimum.

[11] For example, see Mirrlees (1975), Grossman and Hart (1983), Rogerson (1985), Jewitt (1988), Sinclair-Desgagné (1994), Conlon (2009), and Jung and Kim (2015) among others.

[12] Mirrlees (1975) showed, in one-signal cases, i.e., $x \in \mathbb{R}$, $\mu > 0$ when $f(x|a)$ satisfies the monotone likelihood ratio property (MLRP), whereas Holmstrom (1979) showed it when $F(x|a)$ satisfies the first-order stochastic dominance (FOSD) condition.

[13] See Lemma 1 in Jewitt (1988).

Lemma 1 $\mu > 0$.

Observe from (3) that, especially when the principal is risk-neutral, the agent's optimal contract, $s^o(\mathbf{x})$, depends on signal $\mathbf{x}$ only through $\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$. That is, $\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$ becomes a sufficient statistic for $\mathbf{x}$ about a^o for designing $s(\mathbf{x})$, implying that what matters to the risk-neutral principal when she designs a contract for her agent is $\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$ rather than $\mathbf{x}$ itself. In fact, $\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$ is the information about a contained in signal $\mathbf{x}$, indicating that how likely it is that the agent has taken a^o rather than some other nearby action when signal $\mathbf{x}$ is realized. Based on this observation, Jung and Kim (2015) showed that analyzing principal-agent problems directly based on $\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$ generally has an advantage over analyzing them based on signal vector $\mathbf{x}$. Thus, as in Jung and Kim (2015), we derive conditions under which the first-order approach can be justified based on $\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$.

We denote $\tilde{q} \equiv Q_{a^o}(\tilde{\mathbf{x}}) \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$ as the information variable, which implies that random variable $\tilde{q}$ and random vector $\tilde{\mathbf{x}}$ have a functional relationship such that $q = Q_{a^o}(\mathbf{x}) \equiv \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$, where q is the realized value of $\tilde{q}$. Note that q is a function of $\mathbf{x}$ but it also depends on a^o. Thus, since $\tilde{q}$ is defined based on the given a^o, the support of $\tilde{q}$, denoted by $[\underline{q}, \overline{q}]$, may depend on a^o. We also denote $G(q|a)$ as the cumulative distribution function of $\tilde{q}$ given a, i.e.,

$$G(q|a) \equiv Pr\left[Q_{a^o}(\tilde{\mathbf{x}}) \leq q|a\right],$$

and $g(q|a)$ as its probability density function.We assume that for any given a^o, $G(q|a)$ exhibits the FOSD, i.e., $G_a(q|a) \leq 0$ for all (q, a).

Actually, $q = \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$ is what the principal thinks the likelihood ratio is, assuming that the agent chose a^o, and given that the principal observes $\mathbf{x}$. When the principal observes $\mathbf{x}$, she can calculate q, so we can imagine her as actually observing q instead. So the principal expects q to be distributed as $G(q|a^o)$, but it is actually distributed as $G(q|a)$.

Based on q, define

$$w(q) \equiv u'^{-1}\left(\frac{1}{\lambda + \mu q}\right) \quad \text{and} \quad \overline{r}(q) \equiv u(w(q)). \tag{5}$$

We see from (3) that $w(q)$ denotes the optimal contract defined on the q-space when it is not constrained by the limited liability constraint, i.e., $s^o(\mathbf{x}) \geq \underline{s}$, whereas $\overline{r}(q)$ denotes the agent's indirect utility also defined on the q-space in that case. Note from (5) that the functional forms of $w(\cdot)$ and $\overline{r}(\cdot)$ depend only on the functional form of $u(\cdot)$, and Lemma

1 guarantees that both $w(q)$ and $r(q)$ are increasing in q.

Thus, the agent's indirect utility given $s^o(\mathbf{x})$ in (3), i.e., $u(s^o(\mathbf{x}))$, can be written as

$$u(s^o(\mathbf{x})) \equiv r(Q_{a^o}(\mathbf{x})) = r(q) = \begin{cases} \overline{r}(q), & \text{when } q \geq q_c, \\ u(\underline{s}), & \text{when } q < q_c, \end{cases} \tag{6}$$

where $q_c = Q_{a^o}(\mathbf{x}_c) \equiv \frac{f_a(\mathbf{x}_c|a^o)}{f(\mathbf{x}_c|a^o)}$ solves $\frac{1}{u'(\underline{s})} = \lambda + \mu q_c$.[14] Also, the agent's expected monetary utility when he takes an effort a under $s^o(\mathbf{x})$ can be written as

$$\begin{aligned} U(s^o(\cdot), a) &= \int u(s^o(\mathbf{x})) f(\mathbf{x}|a) d\mathbf{x} - a \\ &= \int r(q) g(q|a) dq - a. \end{aligned}$$

It is widely known that the agent's limited liability constraint may or may not be binding at the optimum depending on the characteristic variables, $\{\pi(\mathbf{x}),u(s),f(\mathbf{x}|a),\overline{U},\underline{s}\}$, especially on whether the information variable, $\tilde{q} = \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$, is bounded below or not.[15] If it is the case where the agent's limited liability constraint is binding for some $\mathbf{x}$ with the optimal contract, the participation constraint may not be binding (i.e., $\lambda = 0$), and the agent may enjoy some positive rent at the optimum. Thus, to be more general, we denote the agent's expected utility when he takes a^o given $s^o(\mathbf{x})$ as

$$\begin{aligned} U(s^o(\cdot), a^o) &= \int u(s^o(\mathbf{x})) f(\mathbf{x}|a^o) d\mathbf{x} - a^o \\ &= \int r(q) g(q|a^o) dq - a^o \equiv U^o \geq \overline{U}, \end{aligned} \tag{7}$$

where U^o is determined by $\{\pi(\mathbf{x}),u(s),f(\mathbf{x}|a),\overline{U},\underline{s}\}$. Of course, $U^o = \overline{U}$ if the limited liability constraint is not binding at the optimum (i.e., $\lambda > 0$).

The following Lemma 2 plays a fundamental role in driving our main results.

Lemma 2 *For any given a^o, if*
(L1) *$f(\mathbf{x}|a)$ satisfies that $\frac{g(q|a)}{g(q|a^o)}$ is convex in q for all a, and $\xi(q)$ is a function that satisfies the following (L2), (L3), and (L4) where:*
(L2) $\int \xi(q) g(q|a^o) dq = 0$,

[14] Note that λ and μ are functions of $\underline{s}$ and a^o.

[15] This is associated with what is called the Mirrlees' unpleasant theorem. See Mirrlees (1975) for this issue.

(L3) $\int \xi(q) \cdot q \cdot g(q|a^o)dq = 0$, *and*
(L4) $\xi(q)$ *changes sign twice from negative to positive and then to negative as* q *increases, then, we have*

$$\int \xi(q)g(q|a)dq \leq 0, \quad \forall a.$$

Let S_{arg} be a set of contracts that give the agent U^o in (7) as his expected utility when he chooses a^o, and satisfy the original "argmax" incentive constraint at a^o, and S_f be a set of contracts that give the agent the same expected utility as U^o in (7) when he takes a^o, and satisfy the relaxed incentive constraint at that effort level. That is,

$$S_{arg} \equiv \{s(\mathbf{x})|\ s(\mathbf{x}) \text{ satisfies } U(s(\cdot), a^o) = U^o \text{ and } U(s(\cdot), a) \leq U(s(\cdot), a^o), \forall a\}, \tag{8}$$

and

$$S_f \equiv \{s(\mathbf{x})|\ s(\mathbf{x}) \text{ satisfies } U(s(\cdot), a^o) = U^o \text{ and } U_a(s(\cdot), a^o) = 0\}, \tag{9}$$

where $S_{arg} \subseteq S_f$.[16] Then, using Lemma 2, we obtain the following lemma which is one of our key results.

Lemma 3 *For any given a^o, if*
(1a)[= (L1)] $f(\mathbf{x}|a)$ *satisfies that* $\frac{g(q|a)}{g(q|a^o)}$ *is convex in* q *for all* a, *and*
(2a) *(i.e., double crossing property) there exists a contract* $\hat{s}(\mathbf{x}) \in S_{arg}$ *such that* $\hat{r}(q) \equiv u(\hat{s}(\mathbf{x}))$ *crosses* $r(q) \equiv u(s^o(\mathbf{x}))$ *twice starting from above,*
then using the first-order approach is justified.

What Lemma 3 indicates is the following: Note that, since $s^o(\mathbf{x})$ is the optimal contract obtained from the first-order approach, $s^o(\mathbf{x})$ must be in S_f. Then, to guarantee the validity of $s^o(\mathbf{x})$, we need to show that $s^o(\mathbf{x}) \in S_{arg}$. To do this, we consider another contract $\hat{s}(\mathbf{x}) \in S_{arg}$ such that $\hat{r}(q) \equiv u(\hat{s}(\mathbf{x}))$ crosses $r(q) \equiv u(s^o(\mathbf{x}))$ twice starting from above as drawn in Figure 2a.[17] Then, if one thinks of $r(q) - \hat{r}(q)$ as $\xi(q)$ in Lemma 2, it can be easily seen that $\xi(q) = r(q) - \hat{r}(q)$ satisfies the condition (L4) in Lemma 2. Furthermore, as shown in the proof of Lemma 3, the fact that both $s^o(\mathbf{x})$ and $\hat{s}(\mathbf{x})$ are in S_f implies

$$\int [r(q) - \hat{r}(q)]g(q|a^o)dq = 0, \tag{10}$$

[16] Not that, since both S_{arg} and S_f are defined based on given levels of a^o, they vary as a^o changes.

[17] The fact that $\hat{r}(q)$ crosses $r(q)$ twice starting from above (double-crossing) is equivalent to that $\hat{s}(\mathbf{x})$ crosses $s^o(\mathbf{x})$ twice starting from above if $x \in \mathbb{R}$ and $f(x|a)$ satisfies the MLRP.

indicating that condition (L2) in Lemma 2 is satisfied, and

$$\int [r(q) - \hat{r}(q)] \underbrace{q \cdot g(q|a^o)}_{=g_a(q|a^o)} dq = 0, \tag{11}$$

indicating that condition (L3) in Lemma 2 is satisfied. Thus, based on Lemma 2, we derive that, if condition (1a)[=(L1)] is satisfied, the agent's expected utility under $s^o(\mathbf{x})$ is lower than that under $\hat{s}(\mathbf{x})$ for all a except for a^o as drawn in Figure 2b, i.e.,

$$U(s^o(\cdot), a) - U(\hat{s}(\cdot), a) = \int [r(q) - \hat{r}(q)] g(q|a) dq \leq 0, \quad \forall a. \tag{12}$$

This result can be understood as follows: we know $U(s^o(\cdot), a^o) - U(\hat{s}(\cdot), a^o) = 0$. If $a \neq a^o$, due to the double-crossing in Figure 2a, it is more likely that the distribution $g(q|a)$ puts more weighs on the negative values of $r(q) - \hat{r}(q)$ than $g(q|a^o)$, making its average value negative. Our statistical condition (1a) guarantees that this claim holds.

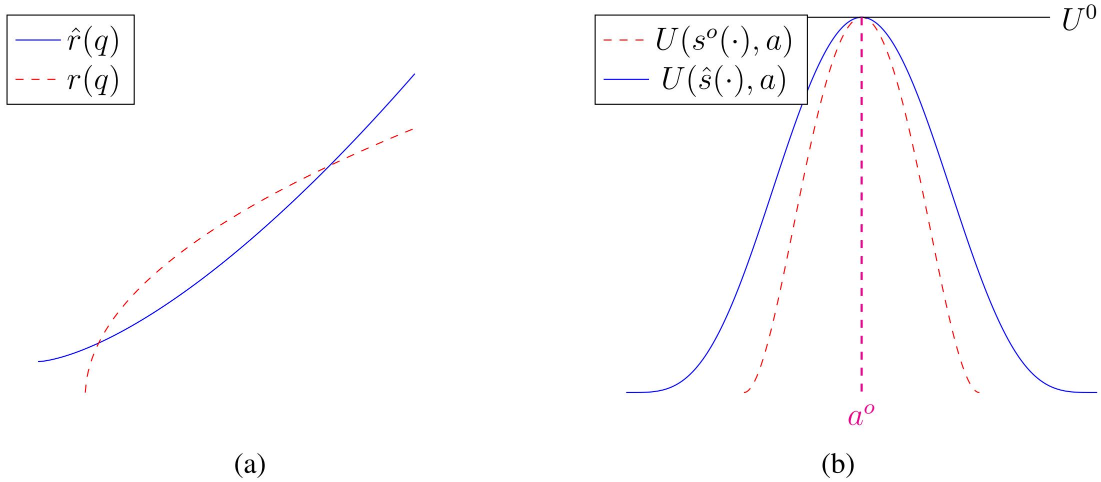


Figure 2: Double Crossing Property

Since $\hat{s}(\mathbf{x}) \in S_{arg}$, we already have

$$U(\hat{s}(\cdot), a) \leq U(\hat{s}(\cdot), a^o) = U^o. \tag{13}$$

Thus, by combining (12) and (13), we have

$$U(s^o(\cdot), a) \leq U(\hat{s}(\cdot), a^o) = U^o, \quad \forall a,$$

which justifies the first-order approach.

Condition (1a) in Lemma 3 is our main statistical condition which is different from the typical statistical conditions in the existing literature such as CDFC (convexity of the distribution function condition) by Mirrlees (1975) and Rogerson (1985), CISP condition (concave increasing set probability condition) by Conlon (2009), and CDFCL (convexity of the distribution function condition for the likelihood ratio) by Jung and Kim (2015). In Section B, we will explain the statistical implication of (1a) as well as the difference between (1a) and the above existing conditions more precisely. However, it is worth to note that (1a) is much easier to verify than the existing conditions in the previous literature and also quite general in that many familiar density functions satisfy it.

For example, in the one-signal case, consider a density function in the exponential family such as

$$f(x|a) = A(a)B(x)e^{\alpha(a)\beta(x)}, \quad x \in \mathbb{R},$$

with $\alpha(a)$ and $\beta(x)$ increasing. Then,

$$\frac{f_a(x|a)}{f(x|a)} = \alpha'(a)\beta(x) + \frac{A'(a)}{A(a)}.$$

Thus, for any given a^o,

$$Q_{a^o}(\tilde{x}) \equiv \frac{f_a(\tilde{x}|a^o)}{f(\tilde{x}|a^o)} = \alpha'(a^o)\beta(\tilde{x}) + \frac{A'(a^o)}{A(a^o)}.$$

Since $\alpha(a)$ and $\beta(x)$ are increasing, $f(x|a)$ satisfies the MLRP. Therefore,

$$G(q|a) = Pr[Q_{a^o}(\tilde{x}) \leq q|a] = Pr[\tilde{x} \leq x|a] = F(x|a),$$

where x solves $\frac{f_a(x|a^o)}{f(x|a^o)} = q$,[18] and

$$g(q|a) = f(x|a)\frac{dx}{dq}.$$

By using $\beta(x) = \left(q - \frac{A'(a^o)}{A(a^o)}\right)\frac{1}{\alpha'(a^o)}$ and $\frac{dx}{dq} = \frac{1}{\alpha'(a^o)\beta'(x)}$, we have

$$g(q|a) = \frac{A(a)}{\alpha'(a^o)}\delta(q)\exp\left\{\frac{\alpha(a)}{\alpha'(a^o)}\left[q - \frac{A'(a^o)}{A(a^o)}\right]\right\},$$

[18]Thus, x which solves $\frac{f_a(x|a^o)}{f(x|a^o)} = q$ is a function of q given a^o, i.e., $x(q; a^o)$.

where $\delta(q) \equiv \frac{B(x)}{\beta'(x)}$.[19] As a result,

$$\frac{g(q|a)}{g(q|a^o)} = \frac{A(a)}{A(a^o)} \exp\left\{\frac{\alpha(a)-\alpha(a^o)}{\alpha'(a^o)}\left[q - \frac{A'(a^o)}{A(a^o)}\right]\right\},$$

which is convex in q for any given a^o and for all a, satisfying condition (1a).

More generally, even in the multi-signal case (i.e., $\mathbf{x} \in \mathbb{R}^n$), condition (1a) holds if $f(\mathbf{x}|a)$ generates, for any given a^o,

$$\frac{g_a(q|a)}{g(q|a)} = A(a)q + D(a),$$

which is the case for most exponential families of density functions (including normal, gamma, etc.) Note that

$$\frac{g(q|a)}{g(q|a^o)} = \exp\left\{\int_{a^o}^{a} \frac{g_a(q|t)}{g(q|t)} dt\right\}.$$

Thus,

$$\frac{g(q|a)}{g(q|a^o)} = \exp\left\{\hat{A}(a)q + \hat{D}(a)\right\},$$

where $\hat{A}(a) \equiv \int_{a^o}^{a} A(t)dt$ and $\hat{D}(a) \equiv \int_{a^o}^{a} D(t)dt$. Therefore, one can easily see that condition (1a) is satisfied.

On the other hand, condition (2a) in Lemma 3 needs to be elaborated more. In general, directly verifying whether condition (2a) is satisfied or not is not easy. In other words, for a given principal-agent problem, finding a proxy contract $\hat{s}(\mathbf{x}) \in S_{arg}$ which satisfies the double crossing property (i.e., condition (2a)), if any, is not straightforward. Therefore, in what follows, we investigate the conditions which sufficiently guarantee condition (2a) and are easier to verify.

4 Verifying the Double Crossing Property

In verifying the existence of $\hat{s}(\mathbf{x}) \in S_{arg}$ which satisfies the double crossing property between $r(q)$ and $\hat{r}(q)$ (i.e., condition (2a) in Lemma 3), the key point is to find an appropriate proxy contract, $\hat{s}(\mathbf{x}) \in S_{arg}$, with which the double crossing property between $r(q)$ and $\hat{r}(q) \equiv u(\hat{s}(\mathbf{x}))$ can be verified as easy as possible. To do this, however, we have to

[19] Since $\beta(x)$ is increasing, we have $x = \beta^{-1}\left(\frac{q - \frac{A'(a^o)}{A(a^o)}}{\alpha'(a^o)}\right) \equiv \Omega(q)$. Then, $\delta(q) \equiv \frac{B(\Omega(q))}{\beta'(\Omega(q))}$.

distinguish the case in which the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum from the case in which that constraint is not binding for all $\mathbf{x}$ at the optimum. This is because $r(q) \equiv u(s^o(\mathbf{x}))$ in (6) has a different functional form depending on whether the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum or not. In other words, the agent's indirect utility given $s^o(\mathbf{x})$, $u(s^o(\mathbf{x})) \equiv r(q)$, must be bounded below by $u(\underline{s})$ for some low values of q (i.e., $q < q_c$) in the case where his limited liability constraint is binding for some $\mathbf{x}$ at the optimum, whereas it is not bounded below by $u(\underline{s})$ (i.e., $r(q) = \overline{r}(q)$, $\forall q$) in the case where that constraint is not binding for all $\mathbf{x}$ at the optimum. This requires a different $\hat{s}(\mathbf{x})$ be introduced to guarantee the double crossing property between $r(q)$ and $\hat{r}(q) \equiv u(\hat{s}(\mathbf{x}))$ in each case.

Associated with the above distinction, it is worth noting that the existing results for the validity of the first-order approach should also be divided into two groups, the results which can be applied only to the case in which the agent's limited liability constraint is not binding for all $\mathbf{x}$ at the optimum, and the results which can be applied even to the case in which the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum.

As mentioned earlier, all the existing results were basically derived to make the agent's expected monetary utility given $s^o(\mathbf{x})$, i.e., $R(a) \equiv \int u(s^o(\mathbf{x}))f(\mathbf{x}|a)d\mathbf{x}$, concave in a. More precisely, the Mirrlees-Rogerson conditions in the one-signal case (i.e., $f(x|a)$, $x \in \mathbb{R}$, should satisfy MLRP and CDFC) were derived based on the fact that $R(a) \equiv \int u(s^o(x))f(x|a)dx$ is concave in a for any "increasing" function $u(s^o(x))$ if $f(x|a)$ satisfies CDFC. Later, other conditions were found to generalize CDFC to the multi-signal case. They include Sinclair-Desgagné's GCDFC (i.e., generalized convexity of the distribution function condition), Conlon's CISP condition (i.e., the concave increasing set probability condition), and Jung and Kim's CDFCL (i.e., convexity of the distribution function condition for the likelihood ratio). All these conditions in the literature contain the property of CDFC because they are basically extended versions of one-signal CDFC to the multi-signal case. Since the results that contain those CDF-type conditions can be used for any "increasing" $u(s^o(\mathbf{x}))$ (or $r(q)$), they can be well applied even to the case in which the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum because $u(s^o(\mathbf{x}))$ (or $r(a)$) is generally increasing in this case. However, they can be applied only to a limited set of cases in that most familiar density functions of signals, $f(\mathbf{x}|a)$, hardly ever satisfy such CDF-type conditions.

To overcome this drawback, Jewitt (1988) proposed another set of conditions in the one-signal case which is not related to the CDF-type conditions, and his conditions were generalized to multi-signal cases by Jung and Kim (2015). Both the condition on $f(\mathbf{x}|a)$

in Jewitt (1988) (i.e., Theorem 1 in Jewitt (1988)) and that in Jung and Kim (2015) (i.e., Proposition 7 in Jung and Kim (2015)) are weaker than the above CDF-type conditions, and thus can be satisfied by some familiar density functions.[20] This is because they were derived to make $R(a) \equiv \int u(s^o(\mathbf{x}))f(\mathbf{x}|a)d\mathbf{x} = \int r(q)g(q|a)dq$ concave in a not for any "increasing" function but for any "increasing and concave" function of $u(s^o(\mathbf{x}))$ (or $r(q)$). However, these conditions, although useful, can be applied only to another limited set of cases in which the agent's limited liability constraint is not binding at the optimum. When the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum, $u(s^o(\mathbf{x}))$ (or $r(q)$) cannot be globally concave because it is equal to the lower bound for a range of values of those $\mathbf{x}$, and then rises. We will elaborate on this issue in Section 4.2, later.

4.1 When the Limited Liability Constraint Never Binds

We first start with the case in which the agent's limited liability constraint is not binding for any $\mathbf{x}$ at the optimum, and provide a set of conditions easier to verify than ((1a),(2a)) in Lemma 3. The agent's limited liability constraint will not be binding for all $\mathbf{x}$ at the optimum if the information variable, $\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)} \in [\underline{q}, \overline{q}]$, has a lower bound, i.e., $\underline{q} > -\infty$, and $\underline{s}$ is low enough such that $\overline{r}(\underline{q})$ in (5) is greater than $u(\underline{s})$.[21] Therefore, to guarantee that the agent's limited liability constraint is not binding for any $\mathbf{x}$ at the optimum as long as the information variable, $\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$, is bounded below, we will assume that that $\underline{s}$ is low enough.[22]

When the agent's limited liability constraint is not binding for all $\mathbf{x}$ at the optimum, (3) reduces to

$$\frac{1}{u'(s^o(\mathbf{x}))} = \lambda + \mu \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)} = \lambda + \mu q, \qquad \text{for all } q, \tag{14}$$

and (6) reduces to

$$u(s^o(\mathbf{x})) \equiv r(Q_{a^o}(\mathbf{x})) = r(q) = \overline{r}(q), \qquad \text{for all } q. \tag{15}$$

Furthermore, since the agent's participation constraint must be binding (i.e., $\lambda > 0$) in this

[20] For the examples of such density functions, see Jewitt (1988, p. 1183)

[21] This is the case to which Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015) can be applied.

[22] Actually, $w(\underline{q}) = u'^{-1}(\frac{1}{\lambda+\mu\underline{q}}) \geq \underline{s}$.

case, (7) also reduces to

$$U(s^o(\cdot), a^o) = \int u(s^o(\mathbf{x}))f(\mathbf{x}|a^o)d\mathbf{x} - a^o = \int r(q)g(q|a^o)dq - a^o \equiv U^o = \overline{U}. \quad (16)$$

Then, based on (14), (15), and (16), we have the following proposition.

Proposition 1 *Given that the limited liability constraint does not bind, if, for any given a^o,*
(1a) $\frac{g(q|a)}{g(q|a^o)}$ *is convex in q for all a.*
(2b) $m(a) \equiv \int qg(q|a)dq$ *is concave in a.*
(3b) $r(q)$ *is concave in q.*
then the first-order approach is justified.

The conditions in Proposition 1 are sufficient to guarantee ((1a),(2a)) in Lemma 3 in cases where the likelihood ratio, $\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$, is bounded below, and the agent's limited liability constraint is not binding at any point at the optimum. Especially, conditions (2b) and (3b) are provided as sufficient conditions for condition (2a) in Lemma 3 (i.e., the double-crossing property between $r(q)$ and $\hat{r}(q)$ for some appropriate proxy contract $\hat{s}(\mathbf{x})$) in this case.

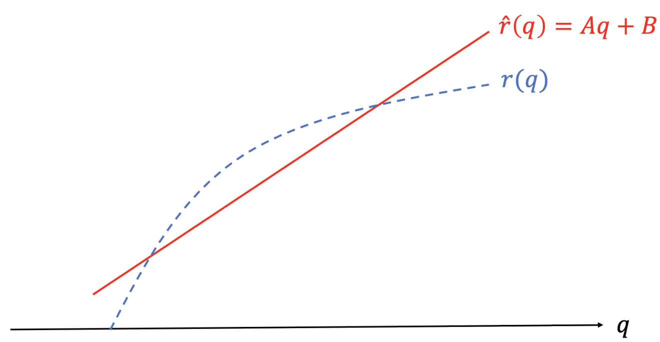


Figure 3: When the Agent's Limited Liability Constraint Does Not Bind

To guarantee the existence of $\hat{r}(q)$ which double-crosses $r(q)$ in this case, we pick a proxy contract, $\hat{s}(\mathbf{x})$, with which the agent's indirect utility is linear in q, i.e., $u(\hat{s}(\mathbf{x})) \equiv \hat{r}(q) = Aq + B$, as shown in Figure 3, where A and B are to be set to satisfy both the participation and the relaxed incentive constraints at a^o. In addition, condition (2b) is given to ensure that the the agent will actually choose a^o given $\hat{s}(\mathbf{x})$, that is, $\hat{s}(\mathbf{x}) \in S_{arg}$. Also, condition (3b) is given to guarantee the double crossing property between $r(q)$ and $\hat{r}(q)$. Thus, one can easily see that

$$U(s^o(\cdot), a) \leq U(\hat{s}(\cdot), a) \leq U(\hat{s}(\cdot), a^o) = U(s^o(\cdot), a^o) = \overline{U}, \quad \forall a,$$

which justifies the first-order approach.

The following Example 1 explains how the conditions in Proposition 1 can be applied to principal-agent problems in which the agent's limited liability constraint is not binding for any $\mathbf{x}$ at the optimum.

Example 1 Consider a one-signal principal-agent problem, $\{\pi(x), u(s), f(x|a), \overline{U}, \underline{s}\}$, where $\pi(x) = x$, $u(s) = \frac{1}{r}s^r$, $r \leq \frac{1}{2}$. We assume that the signal generating function has a simple multiplicative form, $\tilde{x} = h(a)\tilde{\theta}$, where $h(a)$ is increasing with $h(0) = 0$, and $\tilde{\theta}$ is exponentially distributed with mean 1, i.e., the density function of $\tilde{\theta}$ is $p(\theta) = e^{-\theta}$, $\theta \in [0, \infty)$. We also assume that $\underline{s}$ is low enough. Then,

$$f(x|a) = \frac{1}{h(a)} e^{-\frac{x}{h(a)}}, \tag{17}$$

where $E[x|a] = h(a)$. Since, given a^o,

$$\tilde{q} \equiv \frac{f_a(\tilde{x}|a^o)}{f(\tilde{x}|a^o)} = \frac{h'(a^o)}{[h(a^o)]^2}[\tilde{x} - h(a^o)],$$

$f(x|a)$ satisfies MLRP. Since q is bounded below, i.e., $q \geq -\frac{h'(a^o)}{h(a^o)}$, and since $\underline{s}$ is assumed to be low enough, the agent's limited liability constraint is not binding at the optimum. Using $g(q|a)dq = f(x|a)dx$, we derive

$$g(q|a) = \frac{[h(a^o)]^2}{h'(a^o)h(a)} \exp\left(-\frac{1}{h(a)}\left(\frac{[h(a^o)]^2}{h'(a^o)}q + h(a^o)\right)\right). \tag{18}$$

Thus, it can be easily seen that condition (1a) in Proposition 1 holds. Furthermore, from (14) and (15), $r(q) = \frac{1}{r}(\lambda + \mu q)^{\frac{r}{1-r}}$ is concave in q since $r \leq \frac{1}{2}$, which implies that condition (3b) is satisfied. Now, since

$$m(a) \equiv \int qg(q|a)dq = \frac{h'(a^o)}{[h(a^o)]^2}[h(a) - h(a^o)],$$

condition (2b) will hold if $h(a)$ is concave in a. As a result, if $r \leq \frac{1}{2}$ and $h(a) = E[x|a]$ is concave in a, then the first order approach is justified in this case.

Using the first-order approach for the principal-agent problem in the above example can be justified if $h(a)$ is concave in a, when $u(s) = \frac{1}{r}s^r$, $r \leq \frac{1}{2}$ (i.e., concave $r(q)$). First, note that neither $f(x|a)$ in (17) nor $g(q|a)$ in (18) satisfies any of the CDF-type conditions. This indicates that the Mirrlees-Rogerson conditions or any extending conditions of those (i.e., GCDFC of Sinclair-Desgagné (1994), CISP of Conlon (2009), and CDFCL of Jung and Kim (2015)) cannot be used for justifying the first-order approach in this case. However, Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015) are satisfied because $h(a)$ is concave in a and their integral-based conditions are satisfied. Actually, two conditions, (1a) and (2b), in Proposition 1 are sufficient for them.[23] Nevertheless, our conditions in Proposition 1 have advantages that ours is easier to check than those conditions.[24]

Note that if any of the two conditions, (2b) and (3b), in Proposition 1 is violated, not only any of the CDF-type conditions but also Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015) cannot be used for justifying the first-order approach. In Example 1, if $h(a)$ is not concave in a, the CDF-type conditions can never be satisfied because they necessarily requires the concavity of $h(a)$ or $m(a)$. Also, if $r(q)$ is not concave, Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015) cannot be used because they already require the concavity of $r(q)$. However, we will show that our approach can justify the first-order approach even if one of the two conditions is violated. For instance, when condition (3b) is violated, it can be justified by imposing stronger restriction on the concavity of $m(a)$. Conversely, even if condition (2b) is violated, it is also possible by imposing stronger restriction on the concavity of $r(q)$. This is possible because our approach has a flexibility in terms of choosing a proxy contract $\hat{r}(q)$. As a result, we will show that (2b) and (3b) in Proposition 1 have a trade-off relationship, which will be verified through the next two propositions.

Convex $r(q)$ case: violating (3b) in Proposition 1 First, let us consider the case where condition (3b) is violated. For instance, (3b) is not satisfied when $r > \frac{1}{2}$ in Example 1. However, even if the agent's monetary utility in q-space, $r(q)$, is not concave in q (i.e., violating condition (3b) in Proposition 1), the following Proposition 2 shows that the first-order approach can be still justified if we impose stronger restriction on the concavity of

[23]This issue will be dealt with in Section B.1.

[24]In Appendix B, we prove that conditions in Proposition 1 are actually sufficient for Proposition 7 in Jung and Kim (2015).

$m(a)$. To this end, we define $M(a;t)$ as the moment generating function of information variable $\tilde{q}$, i.e.,

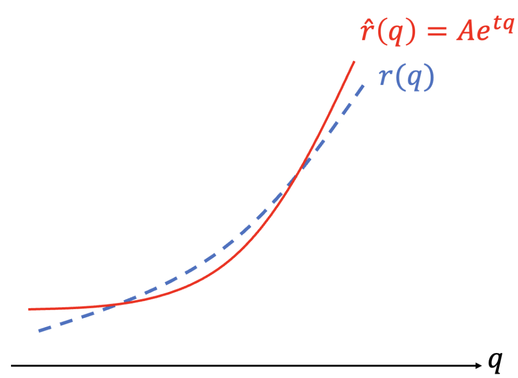
$$M(a;t) \equiv \int e^{tq} g(q|a)dq,$$

which we assume exists with distribution $g(q|a)$ for $t \in T \equiv (\underline{t}, \bar{t}) \subseteq \mathbb{R}$. Note that set T may depend on a^o.

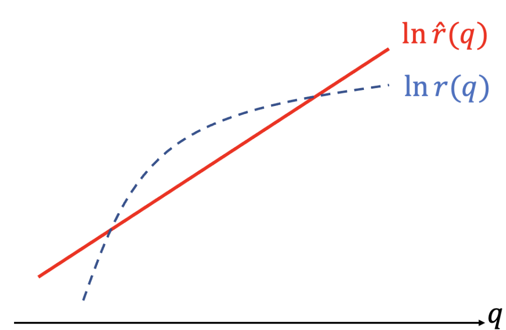
Let us assume that $u(s) > 0$ for all s, implying that $r(q) \equiv u(s^o(\mathbf{x})) > 0$ for all q. To verify the existence of a proxy contract $\hat{r}_t(q) \equiv u(\hat{s}(\mathbf{x}))$, where $\hat{s}(\mathbf{x}) \in S_{arg}$, which double-crosses $r(q)$ in cases where $r(q)$ is possibly convex, we consider a contract $\hat{s}(\mathbf{x})$, with which the agent's indirect utility has an exponential form such as $u(\hat{s}(\mathbf{x})) \equiv \hat{r}_t(q) = A \cdot e^{tq}$, where $A > 0$ and $t > 0$ are to be set to satisfy both the participation and the relaxed incentive constraints as an equality.

Then, we have the following proposition.

Proposition 2 *Given that $u(s) > 0$ for all s and the limited liability constraint does not bind, if, for any given a^o,*
(1a) $\frac{g(q|a)}{g(q|a^o)}$ *is convex in q for all a,*
(2b′) $\phi\left(a; t, \overline{U}\right) \equiv \frac{M_a(a;t)}{M(a;t)} \times [\overline{U} + a]$ *is decreasing in a for any given $t > 0$, and*
(3b′) $\ln r(q)$ *is concave in q,*
then the first-order approach is justified.



(a) Double-Crossing between $\hat{r}(q)$ and $r(q)$



(b) Double-Crossing between $\ln \hat{r}(q)$ and $\ln r(q)$

Figure 4: When $r(q)$ is Convex while $\ln r(q)$ is Concave

As opposed to Proposition 1, where we relied on the double-crossing property between $r(q)$ and $\hat{r}(q)$ which is linear in q as shown in Figure 3, we now construct an *exponential*

$\hat{r}_t(q)$ that double-crosses $r(q)$ starting from above, which is possible even if $r(q)$ is convex, as in Figure 4a. This is possible when $\ln r(q)$, instead of $r(q)$, is concave as we can construct a linear $\ln \hat{r}_t(q)$ that double-crosses $\ln r(q)$ starting from above, as shown in Figure 4b. Thus, condition (3b') guarantees the double-crossing between $\hat{r}_t(q)$ and $r(q)$.

Condition (2b') is needed for $\hat{s}(\mathbf{x}) \in S_{arg}$. Condition (2b') implies that for a given a^o, $\ln M(a;t)$ is concave in $\ln[a+\overline{U}]$ for any $t>0$. Since the proxy contract $\hat{r}_t(q) = Ae^{tq}$ satisfies both the participation and the relaxed incentive constraints, t should satisfy

$$\frac{M_a(a^o;t)}{M(a^o;t)} = \frac{1}{a^o+\overline{U}}, \tag{19}$$

and

$$A = \frac{1}{M_a(a^o;t)}.$$

There exists $t>0$ satisfying (19) for any given a^o, if $\lim_{t\uparrow\bar{t}} \frac{M_a(a^o;t)}{M(a^o;t)} \geq \frac{1}{\overline{U}}$, since $\frac{M_a(a^o;t=0)}{M(a^o;t=0)} = \frac{\int qg(q|a^o)}{1} = 0$ and $0 < \frac{1}{a^o+\overline{U}} \leq \frac{1}{\overline{U}}$ for any a^o.[25] Thus, since

$$\ln[A \cdot M(a;t)] - \ln[a+\overline{U}]$$

is equal to 0 at $a=a^o$ and its derivative at $a=a^o$ is 0 as well, condition (2b') guarantees that $\ln[A \cdot M(a;t)] - \ln[a+\overline{U}]$ has a global maximum at $a=a^o$, implying that $E[\hat{r}_t(q)|a] - [a+\overline{U}] = A \cdot M(a;t) - [a+\overline{U}]$ has a global maximum at $a=a^o$. Consequently, condition (2b') guarantees $\hat{s}(\mathbf{x}) \in S_{arg}$.

It is trivial that the concavity of $\ln r(q)$ (i.e., condition (3b') in Proposition 2) is implied by the concavity of $r(q)$ (i.e., condition (3b) in Proposition 1) because the set of increasing concave functions is closed under composition. Thus, (3b') is weaker than (3b). In turn, condition (2b') in Proposition 2 that for any given $t>0$, $\phi\left(a;t,\overline{U}\right) \equiv \frac{M_a(a;t)}{M(a;t)} \times [\overline{U}+a]$ is decreasing in a is stronger than (2b) in Proposition 1. To see this, note that when $u(s)>0$ for all s, we have $a+\overline{U}>0$ for all a for the existence of the optimal contract satisfying the participation constraint as an equality for any given a^o. Condition (2b') requires that

$$\frac{\phi_a(a;t,\overline{U})}{\phi(a;t,\overline{U})} = \frac{M_{aa}(a;t)}{M_a(a;t)} - \frac{M_a(a;t)}{M(a;t)} + \frac{1}{a+\overline{U}} \leq 0, \quad \text{for all } a \text{ and } t>0,$$

[25] Thus, we assume in Proposition 2 that for any given a^o, $\lim_{t\uparrow\bar{t}} \frac{M_a(a^o;t)}{M(a^o;t)} = \infty$.

which implies that $\lim_{t\downarrow 0} \frac{\phi_a(a;t,\overline{U})}{\phi(a;t,\overline{U})} \leq 0$ for all a. Thus, we have

$$\begin{aligned}
\lim_{t\downarrow 0}\left[\frac{M_{aa}(a;t)}{M_a(a;t)} - \frac{M_a(a;t)}{M(a;t)}\right] &= \frac{M_{aat}(a;t=0)}{M_{at}(a;t=0)} - \frac{M_a(a;t=0)}{M(a;t=0)} \\
&= \frac{m''(a)}{m'(a)} \leq -\frac{1}{a+\overline{U}},
\end{aligned}$$

where the first equality holds by L'Hospital's rule and the second equality holds because $M_t(a;t=0) = m(a)$ and $M_a(a;t=0) = \int g_a(q|a)dq = 0$. Since $a + \overline{U} > 0$ for all a, condition (2b') implies $m''(a) \leq 0$. As (2b') implies (2b), (2b') is stronger than (2b). As a result, Proposition 2 shows that even if $r(q)$ is not concave, the first-order approach can be justified if $m(a)$ is concave enough to satisfy condition (2b'), as long as $\ln r(q)$ is concave. Note that the first-order approach with convex $r(q)$ has never been justified in the literature (e.g., Jewitt (1988)) and Jung and Kim (2015)) unless we impose the CDF-type conditions on the distribution of signals.

The following Example 2 clearly shows the case in which the first order approach can be justified even if $r(q)$ is convex in q.

Example 2 Consider a one-signal principal-agent problem, $\{\pi(x),\ u(s),\ f(x|a),\ \overline{U},\ \underline{s}\}$, where $\pi(x) = x$ and $u(s) = \frac{1}{r}s^r$, $0 < r < 1$. The signal x follows the Poisson distribution with mean $h(a)$ as follows:

$$f(x|a) = \frac{[h(a)]^x}{\Gamma(x+1)} e^{-h(a)},$$

where x is a non-negative integer. Since, given a^o,

$$\tilde{q} \equiv \frac{f_a(\tilde{x}|a^o)}{f(\tilde{x}|a^o)} = \frac{h'(a^o)}{h(a^o)}\tilde{x} - h'(a^o),$$

we have

$$g(q|a) = \frac{[h(a)]^{x(q)}}{\Gamma(x(q)+1)} e^{-h(a)},$$

where $x(q) \equiv \frac{h(a^o)}{h'(a^o)}q + h(a^o)$. Since

$$\frac{g(q|a)}{g(q|a^o)} = \left[\frac{h(a)}{h(a^o)}\right]^{x(q)} e^{h(a^o)-h(a)},$$

we see that condition (1a) is always satisfied. Since the moment generating function of $\tilde{q}$ is given by

$$M(a;t) \equiv \mathbb{E}\left[e^{tq}|a\right] = \exp\left\{\hat{t}\cdot h(a) - t\cdot h'(a^o)\right\},$$

where $\hat{t} \equiv \exp\left[\frac{h'(a^o)}{h(a^o)}t\right] - 1 > 0$, we have

$$\phi\left(a;t,\overline{U}\right) \equiv \frac{M_a(a;t)}{M(a;t)} \times [\overline{U}+a] = \hat{t}\cdot h'(a) \times [\overline{U}+a].$$

Thus, condition (2b') is satisfied if and only if

$$\frac{h''(a)}{h'(a)} \leq -\frac{1}{\overline{U}+a} < 0. \tag{20}$$

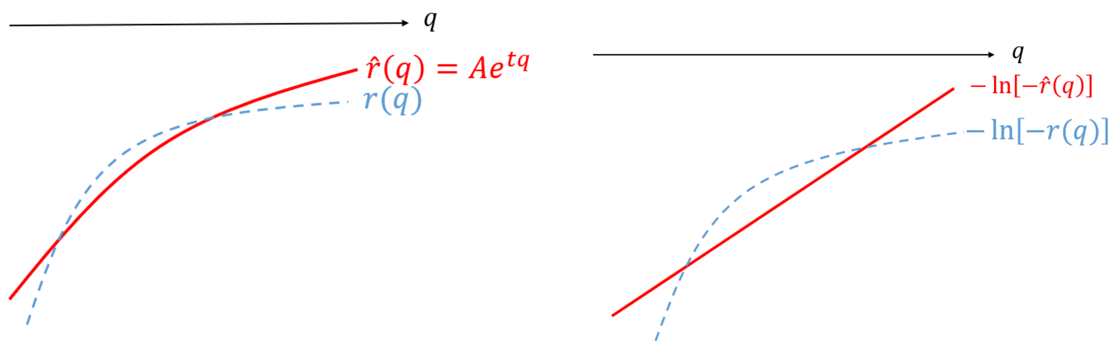
With $r(q) = \frac{1}{r}(\lambda+\mu q)^{\frac{r}{1-r}}$, $\ln r(q) = \frac{r}{1-r}\ln(\lambda+\mu q) - \ln r$ becomes concave in q, satisfying condition (3b'). Thereby with (20) being satisfied, we can justify the first-order approach. Note that (20) implies and therefore is stronger than (2b).

Convex $m(a)$ case: violating (2b) in Proposition 1 Next, let us investigate the case where even if condition (2b) (i.e., the concavity of $m(a)$) is violated, the first-order approach can be justified by requiring a stronger condition than condition (3b). Assume that $u(s) < 0$ for all s, implying that $r(q) < 0$ for any q. To verify the existence of a proxy contract $\hat{r}_t(q) \equiv u(\hat{s}(\mathbf{x}))$, where $\hat{s}(\mathbf{x}) \in S_{arg}$, which double-crosses $r(q)$ which is concave, we consider a contract $\hat{s}_t(\mathbf{x})$, with which the agent's indirect utility has a *negative* exponential form such as $u(\hat{s}_t(\mathbf{x})) \equiv \hat{r}_t(q) = A \cdot e^{tq}$, where $A < 0$ and $t < 0$ are to be set to satisfy both the participation and the relaxed incentive constraints as an equality.

Then, we have the following proposition.

Proposition 3 *Given that $u(s) < 0$, for all s and the limited liability constraint does not binds and for any given a^o, if*
(1a) $\frac{g(q|a)}{g(q|a^o)}$ *is convex in q for all a,*
(2b$''$) $\phi\left(a;t,\overline{U}\right) \equiv \frac{M_a(a;t)}{M(a;t)} \times [\overline{U}+a]$ *is decreasing in a for any given $t < 0$, and*
(3b$''$) $-\ln[-r(q)]$ *is concave in q,*
then the first-order approach is justified.

In contrast to Proposition 2, we here construct a *negative exponential* $\hat{r}_t(q)$ that double-crosses $r(q)$ which is concave, as in Figure 5a. This is possible when $-\ln[-r(q)]$ is concave as in (3b"), as we can construct a linear $-\ln[-\hat{r}_t(q)]$ that double-crosses $-\ln[-r(q)]$



(a) Double-Crossing between $\hat{r}(q)$ and $r(q)$ (b) Between $-\ln[-\hat{r}(q)]$ and $-\ln[-r(q)]$

Figure 5: When $-\ln[-r(q)]$ is Concave

starting from above, as shown in Figure 5b. Condition (2b") guarantees that there exists such a negative exponential proxy contract $\hat{r}_t(q)$ under which the agent would voluntarily participate and choose $a = a^o$, and condition (3b") guarantees double-crossing between $\hat{r}_t(q)$ and $r(q)$, which justifies the first-order approach in this case.

Condition (2b") is sufficient for $\hat{s}(\mathbf{x}) \in S_{arg}$. It implies that for a given a^o, $\ln[M(a;t)]^{-1}$ is concave in $\ln[-(a+\overline{U})]^{-1}$ for any $t < 0$. Since the proxy contract $\hat{r}_t(q) = Ae^{tq}$ satisfies both the participation and the relaxed incentive constraints as an equality, t should satisfy

$$\frac{M_a(a^o;t)}{M(a^o;t)} = \frac{1}{a^o + \overline{U}} < 0, \tag{21}$$

and

$$A = \frac{1}{M_a(a^o;t)} < 0.$$

Here, there exists $t < 0$ satisfying (21) for any given a^o, if we assume $\lim_{t \downarrow \underline{t}} \frac{M_a(a^o;t)}{M(a^o;t)} \leq \frac{1}{\overline{a}+\overline{U}}$, since $\frac{M_a(a^o;t=0)}{M(a^o;t=0)} = 0$ and $\frac{1}{\overline{a}+\overline{U}} \leq \frac{1}{a^o+\overline{U}} < 0$ for any $a^o \in (0, \overline{a}]$. Since

$$\ln[-A \cdot M(a;t)]^{-1} - \ln[-(a+\overline{U})]^{-1}$$

is equal to 0 at $a = a^o$ and its derivative at $a = a^o$ is 0, condition (2b") guarantees that $\ln[-A \cdot M(a;t)]^{-1} - \ln[-(a+\overline{U})]^{-1}$ has a global max at $a = a^o$, implying that $E[\hat{r}_t(q)|a] - [a+\overline{U}] = A \cdot M(a;t) - [a+\overline{U}]$ has a global max at $a = a^o$. Consequently, condition (2b") guarantees $\hat{s}(\mathbf{x}) \in S_{arg}$.

Note that if $-\ln[-r(q)]$ is concave (i.e., (3b") in Proposition 3), $r(q)$ must be concave (i.e., (3b) in Proposition 1) because $-\ln(-y)$ is increasing and convex in $y < 0$. Therefore, (3b”) is a stronger condition than (3b). In turn, (2b”) in Proposition 3 that for any given $t < 0$, $\phi\left(a; t, \overline{U}\right) \equiv \frac{M_a(a;t)}{M(a;t)} \times [\overline{U} + a]$ is decreasing in a turns out to be weaker than (2b) in Proposition 1 under some conditions. To see condition (2b) implies condition (2b”), note that when $u(s) < 0$ for all s, we have $a + \overline{U} < 0$ for all $a \in (0, \overline{a}]$ for the existence of the optimal contract for any given a^o. Condition (2b”) requires that

$$\frac{\phi_a(a; t, \overline{U})}{\phi(a; t, \overline{U})} = \frac{M_{aa}(a; t)}{M_a(a; t)} - \frac{M_a(a; t)}{M(a; t)} + \frac{1}{a + \overline{U}} \leq 0, \quad \text{for all } a \text{ and } t < 0,$$

which implies that $\lim_{t\uparrow 0} \frac{\phi_a(a;t,\overline{U})}{\phi(a;t,\overline{U})} \leq 0$ for any a. Assume that $\frac{\phi_a(a;t,\overline{U})}{\phi(a;t,\overline{U})}$ is increasing in $t < 0$ for any given a. Then, condition (2b”) is equivalent to the condition that $\lim_{t\uparrow 0} \frac{\phi_a(a;t,\overline{U})}{\phi(a;t,\overline{U})} \leq 0$ for any a. Thus, we have

$$\lim_{t\uparrow 0} \frac{\phi_a(a; t, \overline{U})}{\phi(a; t, \overline{U})} = \frac{m''(a)}{m'(a)} + \underbrace{\frac{1}{a + \overline{U}}}_{<0} \leq 0,$$

which is implied by $m''(a) \leq 0$ since $a + \overline{U} < 0$ for all a. Therefore, condition (2b) implies condition (2b”) under the condition that $\frac{\phi_a(a;t,\overline{U})}{\phi(a;t,\overline{U})}$ is increasing in $t < 0$ for any given a.

The following example clearly shows the case in which the first order approach can be justified even if $m(a)$ is convex in a.

Example 3 Consider the same one-signal principal-agent problem, $\{\pi(x), u(s), f(x|a), \overline{U}, \underline{s}\}$, where $\pi(x) = x$ and $u(s) = \frac{1}{r}s^r$, $r < 0$. Note that since $r < 0$, $r(q)$ has negative values and is concave. And, note that in this case, $\overline{U} < 0$ and $a \in (0, \overline{a}]$ where $\overline{a} < -\overline{U}$ for the existence of the optimal contract for any a^o. We assume that the signal generating function has a simple multiplicative form, $\tilde{x} = h(a)\tilde{\theta}$, where $h(a)$ is increasing with $h(0) = 0$ and $\tilde{\theta}$ is exponentially distributed with mean 1, i.e., the density function of $\tilde{\theta}$ is $\phi(\theta) = e^{-\theta}$, $\theta \in [0, \infty)$. We also assume that $\underline{s}$ is low enough. Then,

$$f(x|a) = \frac{1}{h(a)} e^{-\frac{x}{h(a)}},$$

where $E[x|a] = h(a)$. Since given a^o,

$$\tilde{q} \equiv \frac{f_a(\tilde{x}|a^o)}{f(\tilde{x}|a^o)} = \frac{h'(a^o)}{h^2(a^o)}[\tilde{x} - h(a^o)],$$

$f(x|a)$ satisfies MLRP. Since q is bounded below, i.e., $q \geq -\frac{h'(a^o)}{h(a^o)}$, and since $\underline{s}$ is assumed to be low enough, the limited liability is assumed not to bind at the optimum. Using $g(q|a)dq = f(x|a)dx$, we derive

$$g(q|a) = \frac{h^2(a^o)}{h'(a^o)h(a)} \exp\left\{-\frac{h(a^o)}{h(a)}\left[\frac{h(a^o)}{h'(a^o)}q + 1\right]\right\}.$$

Thus, it can be easily seen that condition (1a) in Proposition 3 holds. Its moment generating function is given by[26]

$$M(a;t) = \frac{h^2(a^o)}{h^2(a^o) - t \cdot h'(a^o)h(a)} \exp\left[-t\frac{h'(a^o)}{h(a^o)}\right], \tag{22}$$

which is well defined when $t < 0$. Since from (22)

$$\phi(a;t,\overline{U}) \equiv \frac{M_a(a;t)}{M(a;t)}[\overline{U} + a] = -\frac{h'(a)}{h(a) - \hat{t}}[\overline{U} + a],$$

where $\hat{t} = \frac{h^2(a^o)}{t \cdot h'(a^o)} < 0$, we have

$$\frac{\phi_a(a;t,\overline{U})}{\phi(a;t,\overline{U})} = \frac{h''(a)}{h'(a)} + \frac{1}{\overline{U} + a} - \frac{h'(a)}{h(a) - \hat{t}}.$$

Thus, if

$$\frac{h''(a)}{h'(a)} \leq \underbrace{-\frac{1}{\overline{U} + a}}_{>0}, \quad \forall a \in (0, \overline{a}], \tag{23}$$

which might hold even with convex $h(a)$, condition (2b") is satisfied. With $r(q) = \frac{1}{r}(\lambda + \mu q)^{\frac{r}{1-r}}$, $-\ln[-r(q)] = \ln(-r) - \frac{r}{1-r}\ln(\lambda + \mu q)$ becomes concave in q, satisfying condition (3b"). Thereby with (23) being satisfied, we can justify the first-order approach.

Finally, the value of $\hat{t} = \frac{h^2(a^o)}{t \cdot h'(a^o)}$ satisfying (21) or $\phi(a^o;t,\overline{U}) = \frac{M_a(a^o;t)}{M(a^o;t)}[\overline{U} + a^o] = 1$ is

$$\hat{t} = h(a^o) + h'(a^o)[\overline{U} + a^o].$$

Thus, for $\hat{t} < 0$, we must have

$$h(a) + h'(a)[\overline{U} + a] < 0, \quad \forall a \in (0, \overline{a}].$$

[26]Equation (22) is derived in Appendix B.4.

Taking stock Our Proposition 2 (with Example 2) and Proposition 3 (with Example 3) illustrate how our proxy-contract based approach justifies the use of the first-order approach in cases where the previous literature cannot justify its use, i.e., cases where $m(a)$ is convex in a or $r(q)$ is convex in q. Next, we show how our approach can be applied in a similar way to cases where the agent's limited liability binds at the optimum, the case that the previous literature has overlooked as well.

4.2 When the Information Variable, $\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$, Is Unbounded Below

Our approach in Propositions 1 to 3 can be similarly applied even to the cases where the agent's limited liability constraint binds. Whether the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum or not mainly depends on the density function, $f(\mathbf{x}|a)$, and the subsistence wage level $\underline{s}$. Especially, when the density function of the signals has its information variable, $\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$, not bounded below, i.e., $\tilde{q} \in (-\infty, \overline{q})$,[27] we need to have finite $\underline{s}$, i.e., $\underline{s} > -\infty$, to guarantee the existence of the optimal contract. Then, as shown in (3), the agent's limited liability constraint must be binding for some $\mathbf{x}$ at the optimum.

In this case, in order to verify the existence of a proxy function $\hat{r}(q)$ double-crossing $r(q)$, we introduce a proxy contract $\hat{s}_t(\mathbf{x})$ with which the agent's indirect utility becomes an exponential form, i.e., $u(\hat{s}_t(\mathbf{x})) \equiv \hat{r}_t(q) = A \cdot e^{tq} + u(\underline{s})$, where $A > 0$ and $t > 0$ are to be set to satisfy both the participation and the relaxed incentive constraints, i.e., $\hat{s}_t(\mathbf{x}) \in S_f$. Note that $\hat{s}_t(\mathbf{x})$ always satisfies the agent's limited liability constraint (i.e., $\hat{s}_t(\mathbf{x}) \geq \underline{s}$, $\forall \mathbf{x}$) because for any $A > 0$ and $t > 0$, $u(\hat{s}_t(\mathbf{x})) = Ae^{tq} + u(\underline{s}) > u(\underline{s})$ for all q.

We then have the following Proposition 4.

Proposition 4 *Given that the information variable, $\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$, is unbounded below, if, for any given a^o,*
(1a) $\frac{g(q|a)}{g(q|a^o)}$ *is convex in q for all a.*
(2c) $\phi\left(a; t, \overline{U} - u(\underline{s})\right) \equiv \frac{M_a(a;t)}{M(a;t)} \times [\overline{U} - u(\underline{s}) + a]$ *is decreasing in a for any given $t > 0$.*
(3c) $\ln[\overline{r}(q) - u(\underline{s})]$ *is concave in q for all $q > q_c$, where q_c solves $\overline{r}(q_c) = u(\underline{s})$.*
then the first-order approach is justified.

The conditions in Proposition 4 are sufficient to satisfy conditions ((1a),(2a)) in Lemma 3 in the case where the information variable, $\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$, is unbounded below, and thus

[27] This is the case for many familiar distribution functions of the signals (e.g., normal distribution, gamma distribution, Chi-square distribution, etc.).

the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum. Especially, conditions (2c) and (3c) are given as sufficient conditions for condition (2a) in Lemma 3 (i.e., the double crossing property between $r(q)$ and $\hat{r}_t(q)$) in this case.

In Section 4.1, to guarantee the double crossing property between $r(q)$ and $\hat{r}(q)$ in cases where the agent's limited liability constraint is not binding at the optimum, Proposition 1 requires $r(q)$ in (6) be concave in q (i.e., (3b)), and Proposition 2 or 3 requires $\ln r(q)$ or $\ln[-r(q)]^{-1}$ be concave in q instead (i.e., (3b') in Proposition 2 or (3b") in Proposition 3). In Propositions 1, 2, and 3, conditions on $r(q)$ are equivalent to those on $\overline{r}(q)$ as $r(q) = \overline{r}(q)$ for all q. However, in cases where the limited liability constraint binds for $q < q_c$ as in (6), $r(q)$ cannot be concave around q_c, as seen in Figure 6. Figure 6a illustrates that regardless of whether $\overline{r}(q)$ is convex or concave for $q > q_c$, $r(q)$ cannot be concave around q_c, where the limited liability starts to bind.

Instead, we require that the agent's indirect utility given $s^o(\mathbf{x})$ before constrained by the limited liability constraint, i.e., $\overline{r}(q)$ in (5), should satisfy the condition that $\ln\left[\overline{r}(q) - u(\underline{s})\right]$ is concave in $q > q_c$ (i.e., condition (3c)), and introduce a proxy contract $\hat{s}_t(\mathbf{x})$ such that $u(\hat{s}_t(\mathbf{x})) \equiv \hat{r}_t(q) = Ae^{tq} + u(\underline{s})$, which is convex. Given our exponential-affine proxy contract $\hat{r}_t(q)$, $\ln\left[\hat{r}_t(q) - u(\underline{s})\right] = t \cdot q + \ln A$ becomes linear in $q > q_c$, allowing itself to double-cross the concave $\ln\left[\overline{r}(q) - u(\underline{s})\right]$ from above. In turn, our proxy contract $\hat{r}_t(q)$ must double-cross $r(q)$ from above, as in Figure 6b, satisfying (2a) of Lemma 3. Note that since $\ln\left[\overline{r}(q) - u(\underline{s})\right]$ becomes concave if $\overline{r}(q)$ is concave, (3c) of Proposition 4 is weaker than (3b) of Proposition 1.[28]

Condition (2c) guarantees that $\hat{s}(\mathbf{x}) \in S_{arg}$. Note that condition (2c) implies that for a given a^o and for any $t > 0$, $\ln M(a;t)$ is concave in $\ln[a + \overline{U} - u(\underline{s})]$ which is itself concave in $\ln[a + U^o - u(\underline{s})]$ since $U^o \geq \overline{U}$. Thus, condition (2c) guarantees that for any given a^o, $\ln M(a;t)$ is concave in $\ln[a + U^o - u(\underline{s})]$ for any $t > 0$. Since a proxy contract $\hat{r}_t(q) = Ae^{tq} + u(\underline{s})$ satisfies both the participation and the relaxed incentive constraints, t should satisfy

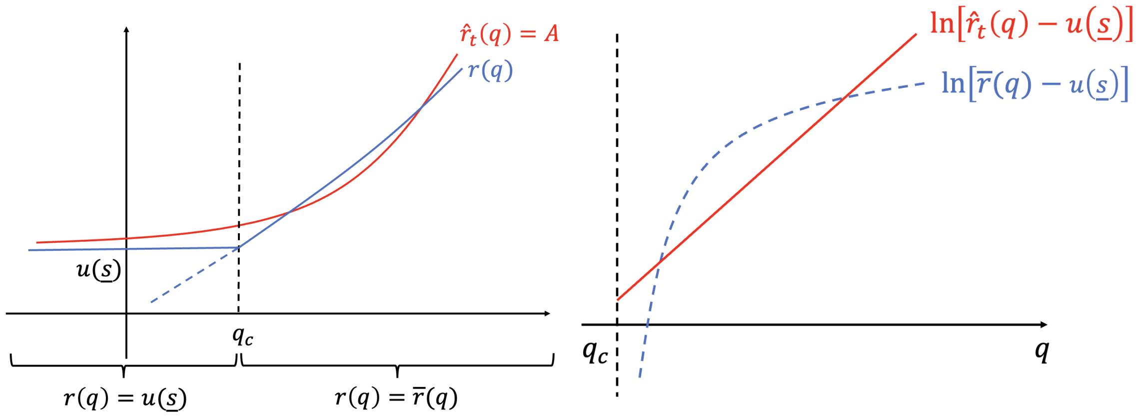
$$\frac{M_a(a^o;t)}{M(a^o;t)} = \frac{1}{a^o + U^o - u(\underline{s})}, \tag{24}$$

and

$$A = \frac{1}{M_a(a^o;t)}.$$

Here, there exists $t > 0$ satisfying (24) for any given a^o, if $\lim_{t \uparrow \bar{t}} \frac{M_a(a^o;t)}{M(a^o;t)} \geq \frac{1}{\overline{U} - u(\underline{s})}$, since

[28] Jewitt (1988) (i.e., Theorem 1) and Jung and Kim (2015) (i.e., Proposition 7) also assume that $\overline{r}(q)$ is concave in their justification of the first-order approach. Our (3c) in Proposition 4 is a weaker condition.



(a) Double-Crossing between $\hat{r}(q)$ and $\overline{r}(q)$ (b) $\ln[\hat{r}_t(q) - u(\underline{s})]$ and $\ln[\overline{r}(q) - u(\underline{s})]$

Figure 6: When the limited liability binds at $q \leq q_c$

$\frac{M_a(a^o;t=0)}{M(a^o;t=0)} = 0$ and $\frac{1}{\overline{U}-u(\underline{s})} \geq \frac{1}{a^o+U^o-u(\underline{s})} > 0$ for any a^o.[29] Since

$$\ln[A \cdot M(a;t)] - \ln[a + U^o - u(\underline{s})]$$

is equal to 0 at $a = a^o$ and its derivative at $a = a^o$ is 0 as well, condition (2c) guarantees that $\ln[A \cdot M(a;t)] - \ln[a + U^o - u(\underline{s})]$ has a global max at $a = a^o$, implying that $E[\hat{r}_t(q)|a] - a - U^o = A \cdot M(a;t) - [a + U^o - u(\underline{s})]$ has a global max at $a = a^o$. Consequently, condition (2c) is sufficient for

$$U(\hat{s}_t(\cdot), a) \equiv \int u(\hat{s}_t(\mathbf{x}))f(\mathbf{x}|a)d\mathbf{x} - a \leq U^o \equiv \int u(s^o(\mathbf{x}))f(\mathbf{x}|a^o)d\mathbf{x} - a^o, \quad \forall a. \tag{25}$$

Both $r(q)$ and $\hat{r}_t(q)$ satisfy the participation and the relaxed incentive constraints, and $\hat{r}_t(q)$ crosses $r(q)$ twice starting from above. Thus, Lemma 3 implies:

$$U(s^o(\cdot), a) \equiv \int r(q)g(q|a)dq - a \leq \int \hat{r}_t(q)g(q|a)dq - a \equiv U(\hat{s}_t(\cdot), a), \quad \forall a. \tag{26}$$

By combining (25) and (26), we finally have

$$U(s^o(\cdot), a) \leq U(\hat{s}_t(\cdot), a) \leq U^o, \ \ \forall a,$$

[29] As in Proposition 2, we also assume for Proposition 4 that for any given a^o, $\lim_{t \uparrow \bar{t}} \frac{M_a(a^o;t)}{M(a^o;t)} = \infty$, which holds generically.

which justifies the first-order approach.

The following two examples illustrate how the conditions in Proposition 4 can be applied to the canonical problems in which the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum (i.e., $\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$ is unbounded below).

Example 4 Consider a one-signal principal-agent problem, $\{\pi(x),\ u(s),\ f(x|a),\ \overline{U},\ \underline{s}\}$, where $\pi(x) = x$, $u(s) = \frac{1}{r}s^r$ with $r < 1$, and $\underline{s} > 0$. Assume that the signal generating function has an additive form such as $\tilde{x} = h(a)+\tilde{\theta}$, $\tilde{\theta} \sim N(0, \sigma^2)$, where $h(a)$ is increasing. Then,

$$f(x|a) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{[x-h(a)]^2}{2\sigma^2}}. \tag{27}$$

Since, given a^o,

$$\tilde{q} \equiv \frac{f_a(\tilde{x}|a^o)}{f(\tilde{x}|a^o)} = \frac{\tilde{x} - h(a^o)}{\sigma^2} h'(a^o),$$

we obtain

$$\tilde{q}|a \sim N\left(\frac{h(a) - h(a^o)}{\sigma^2} h'(a^o), \frac{[h'(a^o)]^2}{\sigma^2}\right).$$

Therefore,

$$g(q|a) = \frac{\sigma}{h'(a^o)\sqrt{2\pi}} \exp\left\{-\frac{\sigma^2}{2[h'(a^o)]^2}\left(q - \frac{h(a) - h(a^o)}{\sigma^2} h'(a^o)\right)^2\right\}. \tag{28}$$

Since $\tilde{x} \in (-\infty, \infty)$, $\tilde{q}$ would be unbounded below, and the agent's limited liability constraint must be binding for some low values of $\tilde{q}$ (i.e., for some low values of $\tilde{x}$).

From (28), we have

$$\frac{g(q|a)}{g(q|a^o)} = \exp\left\{\frac{h(a) - h(a^o)}{h'(a^o)} q - \frac{[h(a) - h(a^o)]^2}{2\sigma^2}\right\},$$

which is convex in q. Therefore, one easily sees condition (1a) is satisfied. Next, note that

$$\frac{d^2}{dq^2} \ln\left[\overline{r}(q) - u(\underline{s})\right] = \frac{\overline{r}'(q)}{\overline{r}(q) - u(\underline{s})}\left(\frac{\overline{r}''(q)}{\overline{r}'(q)} - \frac{\overline{r}'(q)}{\overline{r}(q) - u(\underline{s})}\right). \tag{29}$$

When $r \leq \frac{1}{2}$, we know from (5) that $\overline{r}(q) = \frac{1}{r}(\lambda + \mu q)^{\frac{r}{1-r}}$ is increasing and concave in q, and as $\overline{r}(q) > \overline{r}(q_c) \equiv u(\underline{s})$ for all $q > q_c$, condition (3c) in Proposition 4 is satisfied. Even when $\frac{1}{2} < r < 1$, one can see that $\ln[\overline{r}(q) - u(\underline{s})]$ becomes concave in $q > q_c$ because

$$\begin{aligned}\frac{\overline{r}''(q)}{\overline{r}'(q)} - \frac{\overline{r}'(q)}{\overline{r}(q) - u(\underline{s})} &= \frac{\overline{r}''(q)}{\overline{r}'(q)} - \frac{\overline{r}'(q)}{\overline{r}(q)}\left(1 - \frac{u(\underline{s})}{\overline{r}(q)}\right)^{-1} \\ &\leq \frac{\overline{r}''(q)}{\overline{r}'(q)} - \frac{\overline{r}'(q)}{\overline{r}(q)} = \frac{(2r-1)\mu}{(1-r)(\lambda+\mu q)} - \frac{r\mu}{(1-r)(\lambda+\mu q)} \\ &= -\frac{\mu}{\lambda+\mu q} = -\mu u'(w(q)) < 0,\end{aligned}$$

where the first inequality is from $0 \leq \frac{u(\underline{s})}{\overline{r}(q)} < 1$ for all $q > q_c$ by the fact that $\overline{r}(q) > \overline{r}(q_c) \equiv u(\underline{s}) \geq 0$ for all $q > q_c$. Therefore, regardless of the value of $r < 1$, (3c) in Proposition 4 always holds.

Furthermore, since the moment generating function of $\tilde{q}$ is

$$M(a;t) = \int e^{tq} g(q|a)dq = \exp\left\{t \cdot h'(a^o) \cdot \frac{h(a) - h(a^o)}{\sigma^2} + \frac{[h'(a^o)]^2}{2\sigma^2}t^2\right\} = K{\cdot}\exp\left\{\frac{h'(a^o)}{\sigma^2}t \cdot h(a)\right\},$$

where $K \equiv \exp\left\{\frac{[h'(a^o)]^2}{2\sigma^2}t^2 - \frac{h(a^o)h'(a^o)}{\sigma^2}t\right\}$, we obtain

$$\phi\left(a; t, \overline{U} - u(\underline{s})\right) \equiv \frac{M_a(a;t)}{M(a;t)} \times [\overline{U} - u(\underline{s}) + a] = \frac{h'(a^o)}{\sigma^2}t \cdot h'(a) \times [\overline{U} - u(\underline{s}) + a].$$

Therefore, if

$$\frac{h''(a)}{h'(a)} \leq -\frac{1}{\overline{U} - u(\underline{s}) + a}, \quad \forall a, \tag{30}$$

condition (2c) is satisfied.

This example clearly shows the advantage of the set of conditions in Proposition 4 over the existing sets of conditions. Using the first-order approach in the above example cannot be justified by any of the existing sets of conditions. First, it is easy to see that neither $f(x|a)$ in (27) nor $g(q|a)$ in (28) satisfies any of the the CDF-type conditions. Thus, the Mirrless-Rogerson conditions or any extension of those conditions (e.g., the GCDFC by Sinclair-Desgagné (1994), the CISP condition by Conlon (2009), and the CDFCL by Jung and Kim (2015)) cannot be used for justifying the first-order approach in this case. Furthermore, the Jewitt (1988) conditions and the extension by Jung and Kim (2015) cannot be also used for justifying the first-order approach in this case: both Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015) (i) focus on cases where the agent's limited liability is not binding at the optimum; (ii) require that $r(q)$ be concave in q.[30] In contrast, using the first-order approach in this case can be actually justified by the conditions in Proposition 4 as long as (30) holds, (i) even if the agent's limited liability constraint is binding for $q \leq q_c$; (ii) regardless of whether $\overline{r}(q)$ is concave or convex for $q \geq q_c$.

Actually, the simple signal generating function with a linear form such as $\tilde{x} = h(a)+\tilde{\theta}$, where $\tilde{\theta}$ is normally distributed, has not been able to be used in a wide range of principal-agent problems, because the first-order approach could not be justified under this simple signal generating function by the previous literature, which has been a big obstacle in applying the agency theory to various economic problems. What Example 4 illustrates is that the first-order approach can be validly adopted in the principal-agent problem in which the signals for the agent's effort are normally distributed if one selects $h(a)$ satisfying (30).

Example 5 Consider a one-signal principal-agent problem, $\{\pi(x),\ u(s),\ f(x|a),\ \overline{U},\ \underline{s}\}$, where $\pi(x) = x$, $u(s) = \frac{1}{r}s^r$, $r < 1$, and $\tilde{x} \in (0, \infty)$ follows the gamma distribution with shape parameter $h(a) \geq 0$, i.e.,

$$f(x|a) = \frac{x^{h(a)-1}\beta^{-h(a)}}{\Gamma(h(a))}e^{-\frac{x}{\beta}}. \tag{31}$$

Since, given a^o,

$$\tilde{q} \equiv \frac{f_a(\tilde{x}|a^o)}{f(\tilde{x}|a^o)} = h'(a^o)\ln\tilde{x} - K,$$

where $K \equiv h'(a^o)\ln\beta + \frac{\Gamma'(h(a^o))}{\Gamma(h(a^o))}h'(a^o)$, $\tilde{q}$ is unbounded below. Thus, the agent's limited liability constraint must be binding for some low values of $\tilde{q}$ (i.e., for some low value of $\tilde{x}$). Since the density function of $\tilde{q}$ is given by

$$g(q|a) = \frac{\beta^{-h(a)}}{h'(a^o)\Gamma(h(a))}\exp\left\{\frac{h(a)}{h'(a^o)}(q+K) - \frac{1}{\beta}\exp\left[\frac{q+K}{h'(a^o)}\right]\right\}, \tag{32}$$

we have

$$\frac{g(q|a)}{g(q|a^o)} = \frac{\Gamma(h(a^o))}{\Gamma(h(a))}\beta^{h(a^o)-h(a)}\exp\left\{\frac{h(a)-h(a^o)}{h'(a^o)}(q+K)\right\}, \tag{33}$$

from which one can easily see that condition (1a) in Proposition 4 is satisfied. Also, from Example 4, $\ln\left[\overline{r}(q) - u(\underline{s})\right]$ becomes concave with $\overline{r}(q) = \frac{1}{r}(\lambda+\mu q)^{\frac{r}{1-r}}$ for any $r < 1$, satisfying condition (3c) in Proposition 4.

[30]In those cases where the agent's limited liability constraint is not binding at the optimum, $r(q) = \overline{r}(q)$ for all q as in Section 4.1.

Since the moment generating function of $\tilde{q}$ is given by

$$M(a;t) = \int e^{t \cdot q} g(q|a) dq = e^{-tK} \int x^{t \cdot h'(a^o)} f(x|a) dx = e^{-tK} \times \frac{\Gamma(h(a) + t \cdot h'(a^o))}{\Gamma(h(a))},$$

we obtain

$$\begin{aligned}\phi\left(a; t, \overline{U} - u(\underline{s})\right) &\equiv \frac{M_a(a;t)}{M(a;t)} \times [\overline{U} - u(\underline{s}) + a] \\ &= [\psi(h(a) + t \cdot h'(a^o)) - \psi(h(a))] \times h'(a) \times \left[\overline{U} - u(\underline{s}) + a\right],\end{aligned}$$

where $\psi(z) \equiv \frac{d}{dz} \ln \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$. As the digamma function $\psi(z)$ is increasing and concave in $z > 0$[31], condition (2c) is satisfied if

$$\frac{h''(a)}{h'(a)} \leq -\frac{1}{\overline{U} - u(\underline{s}) + a}, \quad \forall a. \tag{34}$$

Like Example 4, the first-order approach for the principal-agent problem in the above example cannot be justified by any of the existing sets of conditions. One can see that neither $f(x|a)$ in (31) nor $g(q|a)$ in (32) satisfies any of the CDF-type conditions. Thus, the Mirrlees-Rogerson conditions and any extensions of those conditions (e.g., GCDFC by Sinclair-Desgagné (1994), CISP condition by Conlon (2009), and CDFCL by Jung and Kim (2015)) cannot be used for justifying the first-order approach in this case. Furthermore, since, for any given a^o, $\tilde{q} \equiv \frac{f_a(\tilde{x}|a^o)}{f(\tilde{x}|a^o)} = h'(a^o) \ln \tilde{x} - K \in (-\infty, \infty)$, which is unbounded below, the agent's limited liability constraint must be binding for low values of $\tilde{x}$ at the optimum. Thus, by the same reason as in Example 4, Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015) cannot be also used for justifying the first-order approach in this case. However, as shown in Example 5, using the first-order approach in this case can be actually justified by our new conditions in proposition 4 as long as (34) holds.

5 Conclusion

The literature on the first-order approach in the principal-agent problems has been focused on making the agent's expected monetary utility obtained from that approach 'concave' in

[31]For this issue, see e.g., Dragomir, Agarwal and Barnett (2000).

the agent's effort. As relying on such concavity is sometimes overly sufficient, this paper proposes new sets of conditions based on the double-crossing property between a 'proxy' contract and the optimal contract derived from the first-order approach, and shows that our approach can be applied to a wider range of principal-agent problems than the previous literature. Some examples we suggested (i.e., Examples 2, 3, 4, and 5) illustrate interesting cases where the previous literature cannot justify their use of the first-order approach while our new sets of conditions, in contrast, can.

These sets of conditions contain of a statistical condition on the density function of the signals, which is quite general and easy to verify. We also have provided a few alternative sets of conditions (i.e., Propositions 1 to 4) that are derived to be applied specifically to cases in which the agent's limited liability constraint is not binding at the optimum and to other cases in which it is binding for some values of the signal vector at the optimum, respectively. Then, statistical implications of these sets of conditions have been explored, and the comparison between these conditions and the existing conditions has been provided. As our approach extends the applicability of the first-order approach to broader settings than the previous literature allows, including those with the agent's limited liability constraint and convex indirect utility as a function of the likelihood ratio of the signals, we believe that the future research in this area will benefit from this broader applicability of the first-order approach.

References

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

[2] Dragomir, S.S., Agarwal, R.P. and Barnett, N.S. (2000): "Inequalities for Beta and Gamma functions via some classical and new integral inequalities," *Journal of Inequalities and Applications*, 5, 103-165.

[3] Grossman, S. and O. Hart (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, 7-45.

[4] Harris, M. and A. Raviv (1979): "Optimal Incentive Contracts with Imperfect Information," *Journal of Economic Theory*, 20, 231-259.

[5] Holmstrom, B. (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10, 74-91.

[6] Innes, R. D. (1990): "Limited Liability and Incentive Contracting with Ex-ante Action Choices," *Journal of Economic Theory*, 52, 45-67.

[7] Jung, J.Y. and S.K. Kim (2015): "Information Space Conditions for the First-Order Approach in Agency Problems," *Journal of Economic Theory*, 160, 243-279.

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

[9] Jewitt, I., O. Kadan, and J. M. Swinkels (2008): "Moral Hazard with bounded payments," *Journal of Economic Theory*, 143(1), 59-82.

[10] Karlin, S. (1968): *Total Positivity*, Vol. 1. Stanford University Press.

[11] Mirrlees, J. A. (1975): "The Theory of Moral Hazard and Unobservable Behaviour: Part I," mimeo, Nuffield College, Oxford, published in *Review of Economic Studies*, 66, 3-23, 1999.

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

[13] Ross, S. A. (1973): "The Economic Theory of Agency: The Principal's Problem," *American Economic Review*, 63, 134-139.

[14] Shavell, S. (1979): "Risk Sharing and Incentives in the Principal and Agent Relationship," *Bell Journal of Economics*, 10, 55-73.

[15] Sinclair-Desgagné, B. (1994): "The First-Order Approach to Multi-Signal Principal-Agent Problems," *Econometrica*, 62, 459-465.

[16] Spence, M. and R. Zeckhauser (1971): "Insurance, Information, and Individual Action," *American Economic Review*, 61, 380-387.

Appendix A Proofs for Lemmas and Propositions

Proof of Lemma 1. Using (2), the relaxed incentive constraint at a^o is

$$\begin{aligned}
1 &= \int u(s^o(\mathbf{x}))f_a(\mathbf{x}|a^o)d\mathbf{x} \\
&= \int u(s^o(\mathbf{x}))\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}f(\mathbf{x}|a^o)d\mathbf{x} \\
&= \int \left(u(s^o(\mathbf{x})) - E[u(s^o(\mathbf{x}))]\right)\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}f(\mathbf{x}|a^o)d\mathbf{x} \\
&= Cov\left(u(s^o(\mathbf{x})), \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}\right).
\end{aligned}$$

In the above equation, the third equality comes from the fact that $\int E[u(s^o(\mathbf{x}))]f_a(\mathbf{x}|a^o)d\mathbf{x} = 0$ since $E[u(s^o(\mathbf{x}))]$ is constant, and the last equality comes from the fact that $E\left[\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}|a^o\right] = \int f_a(\mathbf{x}|a^o)d\mathbf{x} = 0$. Suppose to the contrary that $\mu \leq 0$. Then, $Cov\left(u(s^o(\mathbf{x})), \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}\right) \leq 0$ from (3), which contradicts $1 > 0$. Therefore, μ must be positive at the optimum.

■

Proof of Lemma 2. Let $\psi(c) \equiv \xi(c)g(c|a^o)$, and define $\psi^{(1)}(q) \equiv \int^q \psi(c)dc = \int^q \xi(c)g(c|a^o)dc$ and $\psi^{(2)}(q) \equiv \int^q \psi^{(1)}(c)dc = \int^q \int^c \xi(t)g(t|a^o)dtdc$. Then, we have $\psi^{(1)}(\overline{q}) = 0$ from (L2). Furthermore, since

$$\begin{aligned}
\int \xi(q)q\cdot g(q|a^o)dq &= q\int^q \xi(c)g(c|a^o)dc\Big|_{\underline{q}}^{\overline{q}} - \int\int^q \xi(c)g(c|a^o)dcdq \\
&= -\int\int^q \xi(c)g(c|a^o)dcdq = 0,
\end{aligned}$$

where the second equality comes from (L2), and the last equality is from (L3), we have $\psi^{(2)}(\overline{q}) = 0$. Note from (L4) that $\psi(c)$ changes sign twice from negative to positive and to negative as c increases. Thus, $\psi^{(1)}(q)$ changes sign once from negative to positive as q increases since $\psi^{(1)}(\overline{q}) = 0$. Since $\psi^{(1)}(q)$ changes sign once from negative to positive as q increases and since $\psi^{(2)}(\overline{q}) = \int \psi^{(1)}(q)dq = 0$, we have

$$\psi^{(2)}(q) = \int^q \int^c \xi(t)g(t|a^o)dtdc \leq 0, \quad \forall q. \tag{A.1}$$

Denote

$$\int \xi(q)g(q|a)dq = \int \xi(q)\frac{g(q|a)}{g(q|a^o)}g(q|a^o)dq = \int \xi(q)\Gamma(q,a)g(q|a^o)dq, \tag{A.2}$$

where $\Gamma(q,a) \equiv \frac{g(q|a)}{g(q|a^o)}$. Then, by taking integration by parts twice, we have

$$\begin{aligned}\int \xi(q)\Gamma(q,a)g(q|a^o)dq &= \Gamma(q,a)\int^{q} \xi(c)g(c|a^o)dc\Big|_{\underline{q}}^{\overline{q}} - \int \left(\int^{q} \xi(c)g(c|a^o)dc\right)\Gamma_q(q,a)dq \\ &= -\Gamma_q(q,a)\int^{q}\int^{c} \xi(t)g(t|a^o)dtdc\Big|_{\underline{q}}^{\overline{q}} + \int \left(\int^{q}\int^{c} \xi(t)g(t|a^o)dtdc\right)\Gamma_{qq}(q,a)dq \\ &= \int \left(\int^{q}\int^{c} \xi(t)g(t|a^o)dtdc\right)\Gamma_{qq}(q,a)dq,\end{aligned} \tag{A.3}$$

where $\Gamma_q(q,a) \equiv \frac{\partial}{\partial q}\Gamma(q,a)$ and $\Gamma_{qq}(q,a) \equiv \frac{\partial^2}{\partial q^2}\Gamma(q,a)$. In (A.3), the second equality comes from the fact that $\psi^{(1)}(\overline{q}) = 0$, and the last equality is from the fact that $\psi^{(2)}(\overline{q}) = 0$. Thus, by using (A.1), (A.2), (A.3), and the fact that $\Gamma_{qq}(q,a) \geq 0$ (i.e., $\frac{g(a|a)}{g(q|a^o)}$ is convex in q), we finally have

$$\int \xi(q)g(q|a)dq = \int \left(\int^{q}\int^{c} \xi(t)g(t|a^o)dtdc\right)\Gamma_{qq}(q,a)dq \leq 0, \quad \forall a.$$

■

Proof of Lemma 3. Since $s^o(\mathbf{x})$ is the optimal contract obtained from the first-order approach, we know $s^o(\mathbf{x}) \in S_f$. Thus, from (9),

$$\begin{aligned}U(s^o(\cdot),a^o) &= \int u(s^o(\mathbf{x}))f(\mathbf{x}|a^o)d\mathbf{x} - a^o \\ &= \int r(q)g(q|a^o)dq - a^o = U^o,\end{aligned} \tag{A.4}$$

and

$$\begin{aligned}U_a(s^o(\cdot),a^o) &= \int u(s^o(\mathbf{x}))f_a(\mathbf{x}|a^o)d\mathbf{x} - 1 \\ &= \int u(s^o(\mathbf{x}))\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}f(\mathbf{x}|a^o)d\mathbf{x} - 1 = \int r(q)\cdot q\cdot g(q|a^o)dq - 1 = 0.\end{aligned} \tag{A.5}$$

Also, since $\hat{s}(\mathbf{x}) \in S_{arg}$,

$$U(\hat{s}(\cdot), a^o) = \int u(\hat{s}(\mathbf{x}))f(\mathbf{x}|a^o)d\mathbf{x} - a^o = \int \hat{r}(q)g(q|a^o)dq - a^o = U^o, \tag{A.6}$$

and

$$U_a(\hat{s}(\cdot), a^o) = \int u(\hat{s}(\mathbf{x}))f_a(\mathbf{x}|a^o)d\mathbf{x} - 1 = \int \hat{r}(q) \cdot q \cdot g(q|a^o)dq - 1 = 0. \tag{A.7}$$

Thus, from (A.4) and (A.6),

$$\int (r(q) - \hat{r}(q))g(q|a^o)dq = 0, \tag{A.8}$$

and, from (A.5) and (A.7),

$$\int (r(q) - \hat{r}(q))q \cdot g(q|a^o)dq = 0. \tag{A.9}$$

By substituting $r(q) - \hat{r}(q)$ for $\xi(q)$ in Lemma 2, one can easily see that $\xi(q) = r(q) - \hat{r}(q)$ satisfies conditions (L2) and (L3) in Lemma 2. Furthermore, as $\hat{r}(q)$ crosses $r(q)$ twice starting from above, $\xi(q) = r(q) - \hat{r}(q)$ satisfies (L4) in Lemma 2. Therefore, given that (L1) is satisfied, we derive from Lemma 2 that

$$\int (r(q) - \hat{r}(q))g(q|a)dq \leq 0, \quad \forall a, \tag{A.10}$$

implying that

$$U(s^o(\cdot), a) \leq U(\hat{s}(\cdot), a), \quad \forall a. \tag{A.11}$$

Since $\hat{s}(\mathbf{x}) \in S_{arg}$,

$$U(\hat{s}(\cdot), a) \leq U(\hat{s}(\cdot), a^o) = U^o, \quad \forall a. \tag{A.12}$$

As a result, by combining (A.11) and (A.12), we have

$$U(s^o(\cdot), a) \leq U(s^o(\cdot), a^o) = U^o, \quad \forall a,$$

which justifies the first-order approach.

■

Proof of Proposition 1. Given a^o, consider an arbitrary contract $\hat{s}(\mathbf{x})$ such that $u(\hat{s}(\mathbf{x})) \equiv$

$\hat{r}(q) = Aq + B$ which satisfies both the participation and the "relaxed" incentive constraints at a^o. Thus, since

$$\hat{R}(a) \equiv \int \hat{r}(q)g(q|a)dq = A \cdot m(a) + B,$$

from the participation constraint in 16), $\hat{r}(q)$ should satisfy

$$\hat{R}(a^o) - a^o = A \cdot m(a^o) + B - a^o = \overline{U}. \tag{A.13}$$

Also, from the relaxed incentive constraint, $\hat{r}(q)$ should satisfy

$$\hat{R}'(a^o) - 1 = A \cdot m'(a^o) - 1 = 0. \tag{A.14}$$

Note that

$$m(a^o) = \int qg(q|a^o)dq = \int \frac{f_a(\mathbf{x}|a^o)}{\cancel{f(\mathbf{x}|a^o)}} \cancel{f(\mathbf{x}|a^o)} d\mathbf{x} = \int f_a(\mathbf{x}|a^o)d\mathbf{x} = 0,$$

which indicates that the expected value of information is always zero (i.e., no information ex-ante).[1] Thus, by solving (A.13) and (A.14), we obtain

$$A = \frac{1}{m'(a^o)}, \ \text{ and } \ B = \overline{U} + a^o. \tag{A.15}$$

Then, using (A.15), we have

$$\hat{R}(a) - a = \frac{1}{m'(a^o)} m(a) + \overline{U} + a^o - a. \tag{A.16}$$

Thus, $\hat{R}(a) - a$ has a maximum value at a^o if

$$\begin{aligned} \hat{R}'(a) - 1 \geq 0 \quad &\Longleftrightarrow \quad \frac{m'(a)}{m'(a^o)} \geq 1, \quad \forall a \leq a^o, \\ \hat{R}'(a) - 1 \leq 0 \quad &\Longleftrightarrow \quad \frac{m'(a)}{m'(a^o)} \leq 1, \quad \forall a > a^o. \end{aligned} \tag{A.17}$$

Since we assume that $g(q|a)$ exhibits the FOSD, i.e., $G_a(q|a) \leq 0$ for all a, we have

[1] Also, note that

$$m'(a^o) = \int qg_a(q|a^o)dq = \int \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)} f_a(\mathbf{x}|a^o)d\mathbf{x} = \int \left(\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}\right)^2 f(\mathbf{x}|a^o)d\mathbf{x} = Var(q|a^o).$$

$m'(a) \geq 0$ for all a. Thus, one can easily see that (A.17) holds if (2b) is satisfied (i.e., $m''(a) \leq 0$). Therefore,

$$\hat{R}(a) - a \leq \hat{R}(a^o) - a^o = \overline{U}, \quad \text{for any given } a^o \text{ and for all } a, \tag{A.18}$$

implying that the proxy contract, $\hat{s}(\mathbf{x})$, also satisfies the original "argmax" incentive constraint, i.e., $\hat{s}(\mathbf{x}) \in S_{arg}$.

Since $r(q) \equiv u(s^o(\mathbf{x}))$ is concave in q by (3b), and since it also satisfies both the participation and the relaxed incentive constraints at a^o, $r(q)$ must cross $\hat{r}(q)$ twice starting from below as drawn in Figure 3. Actually, $r(q)$ and $\hat{r}(q)$ must cross because they both satisfy the same participation constraint at a^o, and they must cross "twice" because $r(q)$ is (increasing and) concave in q whereas $\hat{r}(q)$ is (increasing and) linear in q.[2]

The fact that both $r(q)$ and $\hat{r}(q)$ satisfy the participation constraint at a^o gives

$$\int [r(q) - \hat{r}(q)] g(q|a^o) dq = 0, \tag{A.19}$$

and the fact that they also satisfy the relaxed incentive constraint at a^o gives

$$\int [r(q) - \hat{r}(q)] q \cdot g(q|a^o) dq = 0. \tag{A.20}$$

Thus, by combining (1a) with (A.19), (A.20), and the double crossing property between $r(q)$ and $\hat{r}(q)$, we have from Lemma 2 that

$$\int [r(q) - \hat{r}(q)] g(q|a) dq \leq 0, \quad \forall a. \tag{A.21}$$

Therefore, from (A.18) and (A.21), we finally derive

$$\begin{aligned} U(s^o(\cdot), a) \equiv R(a) - a &\leq \hat{R}(a) - a \\ &\leq \hat{R}(a^o) - a^o = \overline{U} = U(s^o(\cdot), a^o), \quad \text{for any given } a^o \text{ and for all } a, \end{aligned}$$

which justifies the first-order approach.

■

Proof of Proposition 2. For any given a^o, consider a proxy contract $\hat{s}(\mathbf{x})$ such that

[2] $\hat{r}(q)$ crosses $r(q)$ only once, then it is not possible for both $\hat{r}(q)$ and $r(q)$ to induce the same a^o. For more detailed explanation for this, see Innes (1990).

$u(\hat{s}(\mathbf{x})) \equiv \hat{r}_t(q) = Ae^{tq}$, where $A > 0$ and $t > 0$ are to be set to satisfy both the participation constraint and the relaxed incentive constraint at a^o. Since

$$\hat{R}(a) \equiv \int \hat{r}_t(q)g(q|a)dq = AM(a;t),$$

from the participation constraint, we have

$$\hat{R}(a^o) - a^o = A \cdot M(a^o;t) - a^o = \overline{U}, \tag{A.22}$$

and from the relaxed incentive constraint, we have

$$\hat{R}'(a^o) - 1 = A \cdot M_a(a^o;t) - 1 = 0, \tag{A.23}$$

Let (A^o, t^o) be the solution for (A.22) and (A.23). Then,

$$A^o = \frac{1}{M_a(a^o;t^o)}, \tag{A.24}$$

and

$$\frac{M_a(a^o;t^o)}{M(a^o;t^o)} = \frac{1}{\overline{U} + a^o}.\text{[3]} \tag{A.25}$$

Now, to make $\hat{s}(\mathbf{x}) \in S_{arg}$, that is, the agent will choose a^o under $\hat{s}(\mathbf{x})$, we need to show that

$$\hat{R}(a) - a \leq \hat{R}(a^o) - a^o = \overline{U}, \quad \forall a \in [0, \infty),$$

which reduces to

$$\ln M(a;t^o) + \ln A^o \leq \ln[\overline{U} + a], \quad \forall a \in [0, \infty).$$

Define $\Phi(a;a^o) \equiv \ln M(a;t^o) + \ln A^o - \ln[\overline{U} + a]$. Then, from (A.22) and (A.23), we already know that $\Phi(a^o;a^o) = 0$ and $\Phi_a(a^o;a^o) = 0$. Thus, if $\Phi_a(a;a^o) \geq 0, \forall a < a^o$, and $\Phi_a(a;a^o) \leq 0, \forall a > a^o$, the agent will take a^o under $\hat{s}(\mathbf{x})$, i.e., $\hat{s}_t(\mathbf{x}) \in S_{arg}$. Note that

[3] Throughout this paper, we assume for the existence of $t^o > 0$ satisfying this equation (A.25) that for any given a^o, $\lim_{t\to\infty} \frac{M_a(a^o;t)}{M_a(a^o;t)} = \infty$. To see it, note that for any given a^o, we have $\frac{M_a(a^o;t=0)}{M(a^o;t=0)} = \frac{E[q]}{1} = 0$. And, note that since $M_a(a^o;t) = E[qe^{tq}|a^o] = M_t(a^o;t)$ and since $M_{tt}(a^o;t)M(a^o;t) - M_t^2(a^o;t) = E[q^2e^{tq}] \cdot E[e^{tq}] - (E[qe^{tq}])^2 \geq 0$ by Cauchy-Schwarz inequality, $\frac{M_a(a^o;t)}{M(a^o;t)} = \frac{M_t(a^o;t)}{M(a^o;t)}$ is increasing in t. Therefore, that assumption guarantees the unique existence of $t^o > 0$ satisfying (A.25).

$$\begin{aligned}\Phi_a(a;a^o) &= \frac{M_a(a;t^o)}{M(a;t^o)} - \frac{1}{\overline{U}+a} \\ &= \frac{1}{\overline{U}+a} \times \left\{ \frac{M_a(a;t^o)}{M(a;t^o)} \cdot [\overline{U}+a] - 1 \right\} \\ &= \frac{1}{\overline{U}+a} \times \left\{ \phi(a;t^o,\overline{U}) - \phi(a^o;t^o,\overline{U}) \right\},\end{aligned}$$

where $\phi(a;t,\overline{U}) \equiv \frac{M_a(a;t)}{M(a;t)} \cdot [\overline{U}+a]$ and the last equality comes from the fact that $\phi(a^o;t^o,\overline{U}) = 1$ by (A.25). Since $\overline{U}+a > 0$, if $\phi(a;t^o,\overline{U})$ is decreasing in a, then $\hat{s}_t(\mathbf{x}) \in S_{arg}$. Thus, condition (2b') sufficiently guarantees that $\hat{s}_t(\mathbf{x}) \in S_{arg}$.

Since both $\hat{r}_t(q)$ and $r(q)$ satisfy the same participation constraint at a^o, $\hat{r}_t(q)$ and $r(q)$ should be cross. But, they should cross at least more than once because they should also satisfy the same incentive constraint, and so do $\ln \hat{r}_t(q)$ and $\ln r(q)$. Actually, $\ln \hat{r}_t(q)$ crosses $\ln r(q)$ "twice" starting from above since $\ln \hat{r}_t(q)$ is linear in q and $\ln r(q)$ is concave in q, and so does $\hat{r}_t(q)$ against $r(q)$.

Consequently, from Lemma 3, we have

$$R(a) - a \leq \hat{R}(a) - a \leq \hat{R}(a^o) - a^o = R(a^o) - a^o = \overline{U}, \quad \forall a \in [0,\infty),$$

which justifies the first order approach.

■

Proof of Proposition 4. For any given a^o, consider a proxy contract $\hat{s}_t(\mathbf{x})$ such that $u(\hat{s}_t(\mathbf{x})) \equiv \hat{r}_t(q) = Ae^{tq} + u(\underline{s})$, which satisfies both the participation constraint and the "relaxed" incentive constraint at a^o. Therefore, since

$$\hat{R}(a) \equiv \int \hat{r}_t(q) g(q|a) dq = A \cdot M(a;t) + u(\underline{s}),$$

from the participation constraint in (7) and the relaxed incentive constraint, $\hat{r}_t(q)$ should satisfy

$$\hat{R}(a^o) - a^o = A \cdot M(a^o;t) + u(\underline{s}) - a^o = U^o \geq \overline{U}. \tag{A.26}$$

and

$$\hat{R}'(a^o) - 1 = A \cdot M_a(a^o;t) - 1 = 0. \tag{A.27}$$

Then, solving (A.26) and (A.27) gives (A^o, t^o) such that

$$A^o = \frac{1}{M_a(a^o; t^o)}, \tag{A.28}$$

and

$$\frac{M_a(a^o; t^o)}{M(a^o; t^o)} = \frac{1}{U^o - u(\underline{s}) + a^o}. \tag{A.29}$$

Note that since $\hat{r}_t(q) = A^o e^{t^o q} + u(\underline{s}) > u(\underline{s}) = r(q)$ for all $q \leq q_c$, $\hat{r}_t(q)$ and $r(q)$ cannot cross when $q \leq q_c$. Since $\hat{r}_t(q)$ and $r(q)$ satisfy the same participation constraint and the relaxed incentive constraint, respectively, they must cross at least twice on interval $I_c \equiv (q_c, \infty)$, and so do $\ln[r_t(q) - u(\underline{s})]$ and $\ln[r(q) - u(\underline{s})]$ on I_c. Because $\ln[r_t(q) - u(\underline{s})] = t^o q + \ln A^o$ is linear in q whereas $\ln[r(q) - u(\underline{s})]$ is concave in $q > q_c$ by condition (3c), $\ln[r_t(q) - u(\underline{s})]$ double crosses $\ln[r(q) - u(\underline{s})]$ starting from above on interval I_c, and so does $r_t(q)$ against $r(q)$ on I_c.

Now, showing that (2c) implies that under the proxy contract $\hat{r}_t(q)$, the agent will choose $a = a^o$ voluntarily will be enough to prove Proposition 4 with the help of Lemma 3. We need to show:

$$\hat{R}(a) - a \leq \hat{R}(a^o) - a^o = U^o,$$

i.e., $A^o M(a; t^o) \leq U^o - u(\underline{s}) + a$, and equivalently, $\ln M(a; t^o) + \ln A^o \leq \ln[U^o - u(\underline{s}) + a]$. To see this, define a new function $\hat{\Phi}(a; a^o) \equiv \ln M(a; t^o) + \ln A^o - \ln[U^o - u(\underline{s}) + a]$. Note that $\hat{\Phi}(a^o; a^o) = 0$ and $\hat{\Phi}_a(a^o; a^o) = 0$ hold due to (A.26) and (A.29), respectively. Therefore, if $\hat{\Phi}_a(a; a^o) \geq 0$ for all $a < a^o$ and $\hat{\Phi}_a(a; a^o) \leq 0$ for all $a > a^o$, $\hat{\Phi}(a; a^o)$ would have a global maximum at $a = a^o$, implying that the agent chooses a^o under $\hat{r}_t(q)$, i.e., $\hat{s}(\mathbf{x}) \in S_{arg}$. Note that

$$\begin{aligned}
\hat{\Phi}_a(a; a^o) &= \frac{M_a(a; t^o)}{M(a; t^o)} - \frac{1}{U^o - u(\underline{s}) + a} \\
&= \frac{1}{\Delta U^o + a} \left\{ \frac{M_a(a; t^o)}{M(a; t^o)} \cdot [\Delta U^o + a] - 1 \right\} \\
&= \underbrace{\frac{1}{\Delta U^o + a}}_{>0} \left\{ \phi(a; t^o, \Delta U^o) - \phi(a^o; t^o, \Delta U^o) \right\},
\end{aligned}$$

where $\Delta U^o \equiv U^o - u(\underline{s})$ and the last equality holds because $\phi(a^o; t^o, \Delta U^o) = 1$ from (A.29). If $\phi(a; t^o, \Delta U^o)$ is decreasing in a, then $\hat{\Phi}(a; a^o)$ achieves its global maximum at $a = a^o$, which guarantees that $\hat{s}_t(\mathbf{x}) \in S_{arg}$.

Finally, it remains to show that condition (3c) implies $\phi(a; t^o, \Delta U^o)$ decreasing in a. Since $\phi(a; t, \Delta\overline{U})$ decreasing in a for any $t > 0$ by condition (2c) where $\Delta\overline{U} \equiv \overline{U} - u(\underline{s})$, then we have $\phi(a; t^o, \Delta\overline{U})$ decreasing in a, or

$$\frac{\phi_a(a; t^o, \Delta\overline{U})}{\phi(a; t^o, \Delta\overline{U})} = \frac{M_{aa}(a; t^o)}{M_a(a; t^o)} - \frac{M_a(a; t^o)}{M(a; t^o)} + \frac{1}{\Delta\overline{U} + a^o} < 0. \tag{A.30}$$

Since $U^o \geq \overline{U}$, we have from (A.30)

$$\frac{\phi_a(a; t^o, \Delta U^o)}{\phi(a; t^o, \Delta U^o)} = \frac{M_{aa}(a; t^o)}{M_a(a; t^o)} - \frac{M_a(a; t^o)}{M(a; t^o)} + \frac{1}{\Delta U^o + a^o} < 0, \tag{A.31}$$

which is equivalent to $\phi(a; t^o, \Delta U^o)$ decreasing in a.

■

Appendix B The Comparison with the Existing Conditions

As explained, our proxy-contract based conditions in Propositions 1, 2, 3 and 4 are different from the existing conditions in the literature in that they are directly derived to satisfy (4), whereas the existing conditions were derived to make the agent's expected monetary utility given $s^o(\mathbf{x})$, i.e., $R(a) \equiv \int u(s^o(\mathbf{x}))f(\mathbf{x}|a)d\mathbf{x})$ concave in a.

As previously explained, the existing results can be categorized into two groups, the results that contain the CDF-type conditions (i.e., Mirrlees-Rogerson's CDFC, Sinclair-Desgagné's GCDFC, Conlon's CISP condition, and Jung and Kim's CDFCL), and the results that do not contain any CDF-type condition but instead contain a restriction on the agent's utility function (i.e., Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015)). The CDF-type conditions on $f(\mathbf{x}|a)$ in the first group, which were given to make $R(a) \equiv \int u(s^o(\mathbf{x}))f(\mathbf{x}|a)d\mathbf{x}$ concave in a for any "increasing" $u(s^o(\mathbf{x}))$, can be applied even to the case in which the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum because $u(s^o(\mathbf{x}))$ is increasing even in such csae. However, these conditions have limitations in that they are hardly satisfied by most familiar density functions.

On the other hand, the conditions on $f(\mathbf{x}|a)$ that appear in the results of the second group were given to make $R(a) \equiv \int u(s^o(\mathbf{x}))f(\mathbf{x}|a)d\mathbf{x}$ concave in a for any "increasing concave" $u(s^o(\mathbf{x}))$. These conditions must be weaker than the above CDF-type conditions because they require that the concavity of $R(a)$ be satisfied for a smaller set of $u(s)$.[4] However, the results in the second group have another limitation in that they cannot be used for the cases in which the signals' density function has its information variable, $\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$, unbounded below, and thus the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum. Unfortunately, many useful probability density functions (e.g., normal, gamma, etc.) belong to this case. In fact, to generate "increasing concave" $u(s^o(\mathbf{x}))$, the results in the second group contain another restriction on the agent's utility function $u(s)$ such that $u(s^o(\mathbf{x})) \equiv r(q)$ is concave in q. However, this restriction cannot be satisfied in this case due to the agent's binding limited liability constraint.

As a result, it is clear that our conditions in Proposition 4, which can be used for the case in which the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum,[5] have the advantage over the existing results. When the agent's limited liability constraint is binding for some $\mathbf{x}$ at the optimum, among the existing results, only the results that contain

[4]These conditions are satisfied by some familiar density functions (e.g., Chi-square, Poisson etc.).

[5]Note that another advantage of Proposition 4 over the previous literature is that we can have a possibly convex $\overline{r}(q)$ for $q \geq q_c$.

the CDF-type conditions can be used. However, the CDF-type conditions are too restrictive to be satisfied by most familiar density functions of the signals. In contrast, our conditions in Proposition 4 (especially condition (1a)) can be satisfied by the wide range of density functions including the normal and the gamma density functions, as shown in Examples 4 and 5.

On the other hand, the advantage of our conditions in Propositions 1, 2, and 3, which can be applied only to the case in which the agent's limited liability is not binding at all, over the existing results needs to be explained more carefully. This is because the results which do not contain the CDF-type conditions (i.e., Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015)) can still be used for this case.[6] Thus, it will be interesting to compare our conditions in Propositions 1, 2, and 3 with the conditions in Proposition 7 in Jung and Kim (2015) which are up to now the most general conditions in the second group.[7] In Section B, we first investigate the statistical implications of condition (1a) and then, based on these statistical implications, compare our new conditions in Propositions 1, 2, and 3 with Proposition 7 in Jung and Kim (2015). In addition, we illustrate the fundamental differences between our proxy-contract based conditions directly derived to satisfy (4) and the existing conditions derived to make $R(a) \equiv \int u(s^o(\mathbf{x}))f(\mathbf{x}|a)d\mathbf{x}$ concave in a.

Proposition 7 in Jung and Kim (2015) states that, if, for any given a^o,

(1J-1) $\int^z G(q|a)dq$ is convex in a for all z,
(1J-2) $m(a) \equiv \int qg(q|a)dq$ is concave in a,[8] and
(2J) $r(q)(=\overline{r}(q))$ is concave in q,

then the first-order approach is justified.

Notice that the above (2J) is identical to (3b) in Proposition 1, and (1J-2) is identical to (2b) in Proposition 1. Thus, to compare the conditions in Proposition 1 with the conditions in Proposition 7 in Jung and Kim (2015), comparing ((1a),(2b)) with ((1J-1),(1J-2)), given that (2J) (or equivalently (3b)) is satisfied, will be enough. To put forth the conclusion first, ((1a),(2b)) in Proposition 1 are sufficient for ((1J-1),(1J-2)) of Jung and Kim (2015) when $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP. Still even in those cases, Propositions 2 and 3 deal with

[6] We already discussed that (3b') in Proposition 2 is weaker than (3b) in Proposition 1 (which is equivalent to Theorem 1 in Jewitt (1988) and Proposition 7 in Jung and Kim (2015)'s condition on the agent's indirect monetary utility $\overline{r}(q)$). Instead, (2b') in Proposition 2 is a stronger restriction than (2b) in Proposition 1.

[7] Jung and Kim (2015) show that their conditions in Proposition 7 are are more general than the conditions in Theorem 1 in Jewitt (1988).

[8] Condition (1J-2) is generally implied by (1J-1) when z goes to infinity. But the reason we list (1J-2) as a separate condition is because $\int G(q|a)dq$ sometimes may not exist.

cases where the previous literature including Jung and Kim (2015) cannot justify the use of the first-order approach. We introduce the concept of TP_3 as conditions ((1J-1),(1J-2)) are, in general, hard to verify. So our strategy here is as follows:

1. First, we introduce another set of conditions called TP_3-based conditions.
2. We show that conditions (TP_3,(1J-2)) are equivalent to ((1a),(2b)) in Proposition 1.
3. We show that conditions (TP_3,(1J-2)) are sufficient for ((1J-1),(1J-2)) and easier to verify. Therefore, ((1a),(2b)) in Proposition 1 are sufficient as well for ((1J-1),(1J-2)).

The TP_3-based conditions are, for any given a^o,[9] as follows:

(TP_3) $g(q|a)$ is totally positive of degree 3 (i.e., TP_3), and
(1J-2) $m(a) \equiv \int qg(q|a)dq$ is concave in a.[10]

Definition 1 (Total Positivity): A function $f(x,a)$, $f: \mathbb{R}^2 \to \mathbb{R}$, is totally positive of degree n (i.e., TP_n) if, for every $x_1 < x_2 < \cdots < x_n$ and $a_1 < a_2 < \cdots < a_n$,

$$T(f,k) \equiv \begin{vmatrix} f(x_1,a_1) & \cdots & f(x_1,a_k) \\ \vdots & \vdots & \vdots \\ f(x_k,a_1) & \cdots & f(x_k,a_k) \end{vmatrix} \geq 0, \qquad \text{for all } k = 1, 2, \cdots, n.$$[11]

To show that conditions ((1a),(2b)) are more general than the above TP_3-based conditions, we first consider the case in which there is a single signal, i.e., $x \in \mathbb{R}$, and $f(x|a)$ satisfies the MLRP, and thus there is a 1:1 relation between $\tilde{q}$ and $\tilde{x}$. Then, we consider the case in which either there are multiple signals, i.e., $\mathbf{x} \in \mathbb{R}^n$, $n \geq 2$, or $f(x|a)$, $x \in \mathbb{R}$, doest not satisfy the MLRP, and thus there is no 1:1 relation between $\tilde{q}$ and $\tilde{\mathbf{x}}$.

B.1 When $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP

When there is a single signal, i.e., $x \in \mathbb{R}$, and its density function, $f(x|a)$, satisfies the MLRP, there is a 1:1 relation between $\tilde{q}$ and $\tilde{x}$ for any given a^o. Then, as previously shown,

$$G(q|a) = F(x|a) \qquad \text{for all } a,$$

[9] The information variable q is defined for a given a^o, thus we assume that a^o is given.

[10] Since TP_3-based conditions are given to sufficiently guarantee ((1J-1),(1J-2)), the first-order approach can be justified if conditions (TP_3), (1J-2), and (2J) are satisfied.

[11] For a detailed explanation of "total positivity", see Karlin (1968)

where x solves $\frac{f_a(x|a^o)}{f(x|a^o)} = q$. We thus have $g(q|a)Q'_{a^o}(x) = f(x|a)$, where $Q'_{a^o}(x) \equiv \frac{dQ_{a^o}(x)}{dx}$ is independent of a. Thus, we have

$$\frac{g(q|a)}{g(q|a^o)} = \frac{f(x|a)}{f(x|a^o)} \quad \text{and} \quad \frac{g_a(q|a)}{g(q|a)} = \frac{f_a(x|a)}{f(x|a)} \text{ for all } a. \tag{B.1}$$

From (B.1), we obtain an interesting result that, when $x \in \mathbb{R}$ and $f(x|a)$ satisfies the MLRP, (1a) reduces to

(1d) $\frac{f(x|a)}{f(x|a^o)}$ is convex in $q = \frac{f_a(x|a^o)}{f(x|a^o)}$ for any given a^o and for all a.

In other words, if $f(x|a)$ satisfies the MLRP for $x \in \mathbb{R}$, condition (1a) can be replaced by condition (1d), which is much easier to verify because it does not require calculating $g(q|a)$ from $f(x|a)$ explicitly. For instance in Example 5, note that, as $\tilde{x} = a + \tilde{\theta}$, $\tilde{\theta} \sim N(0, \sigma^2)$, $x \in \mathbb{R}$ and $f(x|a)$ satisfies the MLRP. Thus, condition (1a) can be easily verified by using condition (1d) without even calculating $g(q|a)$ from $f(x|a)$. From (27), one can derive that

$$q = \frac{f_a(x|a^o)}{f(x|a^o)} = \frac{x - a^o}{\sigma^2},$$

and

$$\frac{f(x|a)}{f(x|a^o)} = \exp\left(-\frac{1}{2\sigma^2}\left(2(a^o - a)x + a^2 - (a^o)^2\right)\right).$$

Since q is linear in x, and $\frac{f(x|a)}{f(x|a^o)}$ is convex in x for any given a^o, it is easy to confirm that condition (1a) is satisfied.

Lemma 4 *Given two functions, $\phi(x)$ and $\psi(x)$, $\phi : \mathbb{R} \to \mathbb{R}$ and $\psi : \mathbb{R} \to \mathbb{R}$, where $\phi(x)$ is increasing in x,*

$$\begin{vmatrix} 1 & \phi(x_1) & \psi(x_1) \\ 1 & \phi(x_2) & \psi(x_2) \\ 1 & \phi(x_3) & \psi(x_3) \end{vmatrix} \underset{(\leq)}{\geq} 0, \quad \textit{for every } x_1 < x_2 < x_3,$$

if and only if $\psi(x)$ is convex (concave) in $\phi(x)$.

Using Lemma 4, we derive the following Lemma 5 and Corollary 6.

Lemma 5 *Given $x \in \mathbb{R}$, $f(x|a)$ is TP_3 if and only if*
(i) $f(x|a)$ *satisfies MLRP.*
(ii) $\frac{f(x|a)}{f(x|a^o)}$ *is convex in* $q = \frac{f_a(x|a^o)}{f(x|a^o)}$ *for any given a^o and for all a (i.e., condition (1d)).*

Corollary 6 *Given that $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP, $f(x|a)$ becomes TP_3 if and only if $g(q|a)$ is TP_3 for any given a^o.*

One thing to note is that Corollary 1 should not be read as "Given that $x \in \mathbb{R}$, $f(x|a)$ is TP_3 if and only if $g(q|a)$ is TP_3 for any given a^o". The statement, "Given that $x \in \mathbb{R}$, $f(x|a)$ is TP_3 if and only if $g(q|a)$ is TP_3 for any given a^o", is true only when $f(x|a)$ satisfies MLRP. This is because, although, as can be seen from (B.1), MLRP for $f(x|a)$ (i.e., $T(f,2) \geq 0$ for every $x_1 < x_2$ and $a_1 < a_2$) implies MLRP for $g(q|a)$ (i.e., $T(g,2) \geq 0$ for every $q_1 < q_2$ and $a_1 < a_2$ for any given a^o), the converse is not always true.

For instance, consider a class of density functions which is a convex mixture of two probability density functions $p_H(x)$ and $p_L(x)$ such that

$$f(x|a) = \alpha(a)p_H(x) + (1-\alpha(a))p_L(x),$$

where $p_H(x) = -6x^2 + 6x$ and $p_L(x) = 1$ with $x \in [0,1]$, and $\alpha(a) \in [0,1]$ is increasing in a.[12] Then,

$$f(x|a) = \alpha(a)(-6x^2 + 6x - 1) + 1,$$

and

$$\frac{f_a(x|a)}{f(x|a)} = \frac{(-6x^2+6x-1)\alpha'(a)}{1+(-6x^2+6x-1)\alpha(a)}.$$

Thus, it is easy to see that $f(x|a)$ does not satisfy the MLRP.

Define, for any given a^o,

$$\tilde{q} \equiv \frac{f_a(\tilde{x}|a^o)}{f(\tilde{x}|a^o)} = \frac{(-6\tilde{x}^2+6\tilde{x}-1)\alpha'(a^o)}{1+(-6\tilde{x}^2+6\tilde{x}-1)\alpha(a^o)}.$$

Then, $\tilde{q}$ is distributed with support $[\frac{\alpha'(a^o)}{\alpha(a^o)-1}, \frac{\alpha'(a^o)}{\alpha(a^o)+2}]$. Thus,

$$G(q|a) = Pr[\tilde{q} \leq q|a] = Pr\left[-6\tilde{x}^2 + 6\tilde{x} - 1 \leq \frac{q}{\alpha'(a^o) - \alpha(a^o)q}\middle| a\right].$$

If $x_1(q)$ and $x_2(q)$ be two roots of $-6x^2 + 6x - 1 = \frac{q}{\alpha'(a^o)-\alpha(a^o)q}$, where $x_1(q) \leq x_2(q)$,

[12] This example is from Jung and Kim (2015).

then $x_2(q) = 1 - x_1(q)$, where $x_1(q) \in [0, \frac{1}{2}]$ is increasing in q, whereas $x_2(q) \in [\frac{1}{2}, 1]$ is decreasing in q. Therefore,

$$G(q|a) = Pr[\tilde{x} \leq x_1(q)|a] + Pr[\tilde{x} \geq x_2(q)|a].$$

Since $f(x|a)$ is symmetric around $x = \frac{1}{2}$,

$$Pr[\tilde{x} \leq x_1(q)|a] = Pr[\tilde{x} \geq x_2(q)|a].$$

Thus,

$$G(q|a) = 2Pr[\tilde{x} \leq x_1(q)|a] = 2F(x_1(q)|a),$$

and

$$g(q|a) = 2f(x_1(q)|a)x_1'(q).$$

Consequently, we have

$$\frac{g_a(q|a)}{g(q|a)} = \frac{f_a(x_1(q)|a)}{f(x_1(q)|a)}.$$

Since $\frac{f_a(x|a)}{f(x|a)}$ is increasing in $x \in [0, \frac{1}{2}]$, we finally have $\frac{g_a(q|a)}{g(q|a)}$ is increasing in q, indicating that $g(q|a)$ satisfies MLRP.

Based on Lemma 5 and Corollary 6, we observe that, when $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP, the following three statements are equivalent.

(1a) For any given a^o, $\frac{g(q|a)}{g(q|a^o)}$ is convex in $q = \frac{f_a(x|a^o)}{f(x|a^o)} = \frac{g_a(q|a^o)}{g(q|a^o)}$ for all a.

(1d) For any given a^o, $\frac{f(x|a)}{f(x|a^o)}$ is convex in $q = \frac{f_a(x|a^o)}{f(x|a^o)}$ for all a, i.e., $f(x|a)$ is TP_3 given the MLRP for $f(x|a)$.

(1e) For any given a^o, $\frac{g(q|a)}{g(q|a_t)}$ is convex in $\frac{g_a(q|a_t)}{g(q|a_t)}$ for all a and a_t, i.e., $g(q|a)$ is TP_3 given the MLRP for $g(q|a)$.[13]

Thus, given that $x \in \mathbb{R}$ and $f(x|a)$ satisfies the MLRP, condition (1a) indicates that $g(q|a)$ is TP_3 for any given a^o.

Therefore, we have the following proposition.

[13]It is worth noting that (1a) is weaker than (1e) in general. This is because (1a) requires that (1e) hold only for $a_t = a^o$ but not for all a_t. However, (1a) and (1e) become equivalent when $x \in \mathbb{R}$ and $f(x|a)$ satisfies the MLRP.

Proposition 5 *Given that $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP, conditions ((TP_3),(1J-2)) and conditions ((1a),(2b)) of Proposition 1 are equivalent.*

As explained above, (1a) is equivalent to (TP_3) provided that $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP. As condition (2b) and Jung and Kim (2015)'s (1J-2) are the same, conditions ((TP_3),(1J-2)) and conditions ((1a),(2b)) of Proposition 1 are equivalent.

To better understand properties of the conditions ((TP_3),(1J-2)) intuitively, the following characteristics of a density function with TP_3 will be useful.

Lemma 7 *If a density function $f(x|a)$ is TP_3, then, for any increasing and concave function $u(x)$,*

$$u^*(a) \equiv \int u(x)f(x|a)dx \text{ is increasing and concave in } \mu(a) \equiv \int xf(x|a)dx.$$

Lemma 7 shows one of the most interesting characteristics of a density function, $f(x|a)$, with TP_3. When the distribution function satisfying TP_3 is combined with any increasing and concave function, the function's expected value becomes increasing and concave in the distribution's mean value. As explained in Lemma 5 and Corollary 6, given that $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP, condition (1a) is equivalent to that $g(q|a)$ is TP_3 for any given a^o. Thus, when $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP, another interpretation for the conditions in Proposition 1 is possible based on Lemma 7 as follows:

Denote

$$U(s^o(\cdot), a) \equiv \int r(q)g(q|a)dq - a = R(a) - a \equiv \xi(m(a)) - \phi(m(a)),$$

where $m(a) \equiv \int qg(q|a)dq$, i.e., the mean value of the distribution $g(q|a)$. Then, since (1a) in Proposition 1 and (TP_3) are equivalent, $R(a)$ becomes concave in $m(a)$ (i.e., $\xi'' \leq 0$) by Lemma 7.[14] Furthermore, since a is convex in $m(a)$ with concave $m(a)$ (i.e., $\phi'' \geq 0$) by the condition (2b), conditions ((1a),(2b),(3b)) in Proposition 1 ensure that, when $x \in \mathbb{R}$ and $f(x|a)$ satisfies the MLRP, $U(s^o(\cdot), a)$ is concave in $m(a)$, which sufficiently guarantees (4) since $m(a)$ is increasing in a. To justify the first-order approach, all the existing results were derived to make $U(s^o(\cdot), a)$ concave in a (to be more precise, $R(a)$ concave in a). However, the first-order approach can be more generally justified by showing that there exists an increasing function of a, such as $m(a)$, in which $U(s^o(\cdot), a)$ is concave. This is

[14]Due to (3b) in Proposition 1, we assume that $r(q) = \bar{r}(q)$ is increasing and concave for all q.

actually what the conditions in Proposition 1 entails. Finally, our conditions ((1a),(2b)) in Proposition 1 imply ((1J-1),(1J-2)) of Jung and Kim (2015). It is because conditions ((TP_3),(1J-2)), which are equivalent to ((1a),(2b)) in Proposition 1, are sufficient but not necessary for conditions ((1J-1),(1J-2)). As shown in Jewitt (1988), conditions ((1J-1),(1J-2)) are necessary and sufficient for $R(a) \equiv \int r(q)g(q|a)dq$ to be increasing and concave in a for any increasing concave function $r(q)$. On the other hand, conditions ((TP_3),(1J-2)) are sufficient but not necessary for $R(a) \equiv \int r(q)g(q|a)dq$ to be increasing and concave in a for any increasing concave function $r(q)$. The sufficient part comes from that $R(a) \equiv \int r(q)g(q|a)dq \equiv \xi(m(a))$ is increasing concave in $m(a)$ by (TP_3), and $m(a)$ is concave in a by (1J-2). However, even if $R(a) \equiv \int r(q)g(q|a)dq \equiv \xi(m(a))$ is increasing concave in a for any increasing concave function $r(q)$ given that $m(a) \equiv \int qg(q|a)dq$ is concave in a (i.e., (1J-1) and (1J-2)), it does not necessarily mean that $R(a)$ is increasing concave in $m(a)$, i.e., $\xi'' \leq 0$. Furthermore, even if $R(a)$ is increasing and concave in $m(a)$ for any increasing and concave function $r(q)$, $g(q|a)$ may not be always TP_3.

In sum, our conditions ((1a),(2b)) in Proposition 1 are sufficient (and therefore stronger than) for Jung and Kim (2015)'s ((1J-1), (1J-2)), when $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP. Still Our Propositions 2 and 3 deal with cases where the previous literature does not justify the use of the first-order approach, as discussed.

B.2 When $\mathbf{x} \in \mathbb{R}^n$, $n \geq 2$, or $f(x|a)$, $x \in \mathbb{R}$, does not satisfy MLRP

When there are multiple signals, i.e., $\mathbf{x} \in \mathbb{R}^n$, $n \geq 2$, or density function $f(x|a)$ does not satisfy the MLRP even if $x \in \mathbb{R}$, there is no 1:1 relation between $\tilde{q}$ and $\tilde{\mathbf{x}}$. Thus, some of the results that are derived when there is a 1:1 relation between $\tilde{q}$ and $\tilde{x}$ in the previous subsection may not hold. However, the main result in the previous subsection still holds even in this case. That is, although there is no relation of inclusion between our conditions ((1a),(2b)) in Proposition 1 and conditions ((1J-1),(1J-2)) of Jung and Kim (2015), the conditions ((1a),(2b)) are more general than the TP_3-based conditions, ((TP_3),(1J-2)). Nevertheless, it is worth noting that there are two non-trivial differences in this case compared with the previous case.

Consider a multi-signal case where there is a random vector $\tilde{\mathbf{x}} = (\tilde{x}_1, \cdots, \tilde{x}_n)$, $n \geq 2$, with density $f(\mathbf{x}|a)$. Although there might be multiple $\mathbf{x}$ satisfying $\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)} = q$, for any given q, one can calculate $g(q|a)$ from $f(\mathbf{x}|a)$ by using the transformation method of random variables.[15] In order to use the transformation method, we introduce another random vector $\tilde{\mathbf{y}} = (\tilde{y}_1, \cdots, \tilde{y}_n)$ with a density function $\hat{f}(\mathbf{y}|a)$ such that $\tilde{y}_j = \tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)}$ and $\tilde{y}_i = \tilde{x}_i$ for all $i = 1, \cdots, n$, $i \neq j$. If there exists a coordinate x_j in which $\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$ is increasing for any $\mathbf{x}_{-j}$, the density function of $\tilde{\mathbf{y}}$ can be expressed as

$$\hat{f}(\mathbf{y}|a) = f(\mathbf{x}|a) \cdot |J| = f(x_j(q, \mathbf{x}_{-j}), \mathbf{x}_{-j}|a) \times \left| \frac{\partial x_j(q, \mathbf{x}_{-j})}{\partial q} \right|,$$

where J is the transformation's Jacobian and $x_j(q, \mathbf{x}_{-j})$ solves $\frac{f_a(x_j, \mathbf{x}_{-j}|a^o)}{f(x_j, \mathbf{x}_{-j}|a^o)} = q$ for given $\mathbf{x}_{-j}$.[16] Note that, in this case, $|J| = \left| \frac{\partial x_j(q, \mathbf{x}_{-j})}{\partial q} \right|$, which is independent of a because $x_j(q, \mathbf{x}_{-j})$ is independent of a. Then, we have:

$$g(q|a) = \int \hat{f}(q, \mathbf{y}_{-j}|a) d\mathbf{y}_{-j} = \int f(x_j(q, \mathbf{x}_{-j}), \mathbf{x}_{-j}|a) \cdot |J| d\mathbf{x}_{-j} = \int_{\mathbf{x} \in X(q)} f(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j},$$

where $X(q) \equiv \left\{ \mathbf{x} | \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)} = q \right\}$. Thus, although

$$\frac{g_a(q|a^o)}{g(q|a^o)} = \frac{\int_{\mathbf{x} \in X(q)} f_a(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}}{\int_{\mathbf{x} \in X(q)} f(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}} = \frac{\int_{\mathbf{x} \in X(q)} q f(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}}{\int_{\mathbf{x} \in X(q)} f(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}} = q = \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$$

for any given a^o even in this case, it is generally true that

$$\frac{g(q|a)}{g(q|a^o)} = \frac{\int_{\mathbf{x} \in X(q)} f(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}}{\int_{\mathbf{x} \in X(q)} f(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}} \neq \frac{f(\mathbf{x}|a)}{f(\mathbf{x}|a^o)}, \text{ and } \frac{g_a(q|a)}{g(q|a)} = \frac{\int_{\mathbf{x} \in X(q)} f_a(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}}{\int_{\mathbf{x} \in X(q)} f(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}} \neq \frac{f_a(\mathbf{x}|a)}{f(\mathbf{x}|a)}, \tag{B.2}$$

where $q = \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$.

[15] In the one-signal case in which $f(x|a)$, $x \in \mathbb{R}$, does not satisfy the MLRP, one easily obtains $g(q|a) = \sum_k f(x_k(q)|a) \left| \frac{dx_k(q)}{dq} \right|$ where $x_k(q) \in X(q) \equiv \left\{ x | \frac{f_a(x|a^o)}{f(x|a^o)} = q \right\}$.

[16] Even in the multi-signal cases where there is no x_j in which $\frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$ is increasing, we observe equation (B.2) still holds. For example, if the support $\{\mathbf{x}|f(\mathbf{x}|a) > 0\}$ can be decomposed into subsets $X_1, \cdots, X_m$ such that $\tilde{\mathbf{y}}$ is a 1:1 transformation of X_k onto one subset of the support $\left\{ \mathbf{y}|\hat{f}(\mathbf{y}|a) > 0 \right\}$, the density function of $\tilde{\mathbf{y}}$ can be expressed by $\hat{f}(\mathbf{y}|a) = \sum_k f(x_j^k(q, \mathbf{x}_{-j}), \mathbf{x}_{-j}|a) \cdot |J_k|$, where $x_j^k(q, \mathbf{x}_{-j})$ solves $\frac{f_a(x_j, \mathbf{x}_{-j}|a^o)}{f(x_j, \mathbf{x}_{-j}|a^o)} = q$ on X_k for given $\mathbf{x}_{-j}$ and J_k is the Jacobian of the transformation on X_k, from which one can easily obtain $g(q|a) = \int \sum_k f(x_j^k(q, \mathbf{x}_{-j}), \mathbf{x}_{-j}|a) \cdot |J_k| d\mathbf{x}_{-j}$.

For instance, consider multi-signal cases where $\tilde{\mathbf{x}} \sim N(\mu(a), \Sigma)$ where $\mu(a) = [\mu_1(a), \cdots, \mu_n(a)]'$ and Σ is a covariance matrix. Then,

$$f(\mathbf{x}|a) = (2\pi)^{-\frac{n}{2}}|\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}[\mathbf{x} - \mu(a)]'\Sigma^{-1}[\mathbf{x} - \mu(a)]\right),$$

and

$$\frac{f_a(\mathbf{x}|a)}{f(\mathbf{x}|a)} = [\mu'(a)]'\Sigma^{-1}[\mathbf{x} - \mu(a)].$$

Thus, we have

$$\tilde{q} \equiv \frac{f_a(\tilde{\mathbf{x}}|a^o)}{f(\tilde{\mathbf{x}}|a^o)} \sim N\left(m(a), \sigma_q^2\right),$$

where $m(a) \equiv [\mu'(a^o)]'\Sigma^{-1}[\mu(a) - \mu(a^o)]$ (thereby $m(a^o) = 0$) and $\sigma_q^2 \equiv [\mu'(a^o)]'\Sigma^{-1}\mu'(a^o)$, and

$$\frac{g_a(q|a)}{g(q|a)} = \frac{q - [\mu'(a^o)]'\Sigma^{-1}[\mu(a) - \mu(a^o)]}{[\mu'(a^o)]'\Sigma^{-1}\mu'(a^o)}[\mu'(a^o)]'\Sigma^{-1}\mu'(a).$$

Therefore, we see: $\frac{g(q|a)}{g(q|a^o)} \neq \frac{f(\mathbf{x}|a)}{f(\mathbf{x}|a^o)}$ and $\frac{g_a(q|a)}{g(q|a)} \neq \frac{f_a(\mathbf{x}|a)}{f(\mathbf{x}|a)}$ for $\forall a \neq a^o$ where $q = [\mu'(a^o)]'\Sigma^{-1}[\mathbf{x} - \mu(a^o)]$. This shows that, when there is no 1:1 relation between $\tilde{\mathbf{x}}$ and $\tilde{q}$, (B.2) is generally true.

From (B.2), one observes two non-trivial differences between this case and the previous Section B.1 where $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP. First, (1a) cannot be reduced to (1d) as in Section B.1. Therefore, to verify condition (1a) in this case, one should explicitly calculate $g(q|a)$ from $f(\mathbf{x}|a)$. Second, condition (1a) is not equivalent to that $g(q|a)$ is TP_3 for any given a^o. In fact, the condition that $g(q|a)$ is TP_3 for any given a^o is stronger than condition (1a) in this case. In order for $g(q|a)$ to be TP_3 for any given a^o, it is needed that, for any given a^o,

(1) $g(q|a)$ satisfies MLRP, and
(2)(=(1e)) $\frac{g(q|a)}{g(q|a_t)}$ is convex in $\frac{g_a(q|a_t)}{g(q|a_t)}$ for all a, a_t.

However, condition (1a) requires neither that $g(q|a)$ satisfy MLRP nor that condition (1e) hold for all a_t.[17] Thus, in contrast with the previous case, when there is no 1:1 relation between $\tilde{q}$ and $\tilde{\mathbf{x}}$, condition (1a) becomes more general than (TP_3). As a result, our conditions ((1a),(2b)) are even more general than conditions ((TP_3),(1J-2)) in this case.

However, there is a meaningful exception even in this case. Consider a density function

[17]Note: (1a) condition mandates that $\frac{g(q|a)}{g(q|a_0)}$ is convex in $q = \frac{g_a(q|a_0)}{g(q|a_0)}$ for all a.

$f(\mathbf{x}|a)$, $\mathbf{x} \in \mathbb{R}^n$, $n \geq 2$, which generates, for any given a^o,

$$\frac{f_a(\mathbf{x}|a)}{f(\mathbf{x}|a)} = \alpha(a) \cdot q + \beta(a), \tag{B.3}$$

where $q = \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)}$, $\alpha(a^o) = 1$, $\beta(a^o) = 0$, and $\alpha(a) \geq 0$. Note that most exponential-family density functions with an appropriate parameterization satisfy (B.3).

Since

$$\frac{f(\mathbf{x}|a)}{f(\mathbf{x}|a^o)} = \exp\left\{\int_{a^o}^{a} \frac{f_a(\mathbf{x}|t)}{f(\mathbf{x}|t)} dt\right\} = \exp\left\{\left(A(a) - A(a^o)\right) q + B(a) - B(a^o)\right\},$$

where $A(a) \equiv \int_0^a \alpha(t)dt$ and $B(a) \equiv \int_0^a \beta(t)dt$, one can see that, for any given q, $\frac{f(\mathbf{x}|a)}{f(\mathbf{x}|a^o)}$ has the same value for all $\mathbf{x} \in X(q) \equiv \left\{\mathbf{x} | \frac{f_a(\mathbf{x}|a^o)}{f(\mathbf{x}|a^o)} = q\right\}$. Thus,

$$\begin{aligned}
\frac{g(q|a)}{g(q|a^o)} &= \frac{\displaystyle\int_{\mathbf{x}\in X(q)} f(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}}{\displaystyle\int_{\mathbf{x}\in X(q)} f(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}} = \frac{\displaystyle\int_{\mathbf{x}\in X(q)} \frac{f(\mathbf{x}|a)}{f(\mathbf{x}|a^o)} f(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}}{\displaystyle\int_{\mathbf{x}\in X(q)} f(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}} \\
&= \frac{\exp\left(\left(A(a) - A(a^o)\right) q + B(a) - B(a^o)\right) \cancel{\displaystyle\int_{\mathbf{x}\in X(q)} f(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}}}{\cancel{\displaystyle\int_{\mathbf{x}\in X(q)} f(\mathbf{x}|a^o) \cdot |J| d\mathbf{x}_{-j}}} = \frac{f(\mathbf{x}|a)}{f(\mathbf{x}|a^o)}.
\end{aligned}$$

Furthermore, from (B.3), we also have

$$\begin{aligned}
\frac{g_a(q|a)}{g(q|a)} &= \frac{\displaystyle\int_{\mathbf{x}\in X(q)} f_a(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}}{\displaystyle\int_{\mathbf{x}\in X(q)} f(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}} = \frac{\displaystyle\int_{\mathbf{x}\in X(q)} \frac{f_a(\mathbf{x}|a)}{f(\mathbf{x}|a)} f(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}}{\displaystyle\int_{\mathbf{x}\in X(q)} f(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}} \\
&= \frac{(\alpha(a)q + \beta(a)) \cancel{\displaystyle\int_{\mathbf{x}\in X(q)} f(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}}}{\cancel{\displaystyle\int_{\mathbf{x}\in X(q)} f(\mathbf{x}|a) \cdot |J| d\mathbf{x}_{-j}}} = \frac{f_a(\mathbf{x}|a)}{f(\mathbf{x}|a)}.
\end{aligned} \tag{B.4}$$

Therefore, even if there is no 1:1 relation between $\tilde{q}$ and $\tilde{\mathbf{x}}$, all the results in Section B.1 equally hold in this special case. That is, condition (1a) reduces to the condition (1d), and verifying (1a) can be replaced by verifying (1d) which does not require to calculate $g(q|a)$

from $f(\mathbf{x}|a)$ directly. Furthermore, because $\alpha(a) \geq 0$, $g(q|a)$ satisfies MLRP from (B.4). Therefore, based on (B.3) and (B.4), one can also see that condition (1a) is equivalent to that $g(q|a)$ is TP_3 for any given a^o.

MLRP for $f(\mathbf{x}|a)$ and $g(q|a)$ When there is no 1:1 relation between $\tilde{q}$ and $\tilde{\mathbf{x}}$, MLRP for $f(\mathbf{x}|a)$ does not always imply the MLRP for $g(q|a)$ for a given a^o. For example, consider a case of discrete signals with $f(x_1, x_2|a) = e^{2\sqrt{a}x_1 + ax_2 - K(a)}$, $x_i \in \{0,1\}$, $i = 1, 2$, $a > 0$, where $K(a) = \log[(1+e^{2\sqrt{a}})(1+e^a)]$. $\frac{f_a(x_1,x_2|a)}{f(x_1,x_2|a)} = \frac{x_1}{\sqrt{a}} + x_2 - K'(a)$ implies that $f(x_1, x_2|a)$ satisfies MLRP. Define $\tilde{q} = \frac{\tilde{x}_1}{\sqrt{a^o}} + \tilde{x}_2 - K'(a^o)$ and let $a^o < 1$. Then, $g(q_1|a) = f(0,0|a) = e^{-K(a)}$, $g(q_2|a) = f(0,1|a) = e^{a-K(a)}$, $g(q_3|a) = f(1,0|a) = e^{2\sqrt{a}-K(a)}$ and $g(q_4|a) = f(1,1|a) = e^{2\sqrt{a}+a-K(a)}$, where $q_1 = -K'(a^o)$, $q_2 = 1 - K'(a^o)$, $q_3 = \frac{1}{\sqrt{a^o}} - K'(a^o)$, and $q_4 = 1 + \frac{1}{\sqrt{a^o}} - K'(a^o)$ so $q_1 < q_2 < q_3 < q_4$. Thus, $\frac{g_a(q_1|a)}{g(q_1|a)} = -K'(a)$, $\frac{g_a(q_2|a)}{g(q_2|a)} = 1 - K'(a)$, $\frac{g_a(q_3|a)}{g(q_3|a)} = \frac{1}{\sqrt{a}} - K'(a)$, and $\frac{g_a(q_4|a)}{g(q_4|a)} = \frac{1}{\sqrt{a}} + 1 - K'(a)$. Since, when $a > 1$, $\frac{g_a(q_2|a)}{g(q_2|a)} = 1 - K'(a) > \frac{1}{\sqrt{a}} - K'(a) = \frac{g_a(q_3|a)}{g(q_3|a)}$, we observe $g(q|a)$ does not satisfy MLRP in general.

B.3 Proof of Appendix B

Proof of Lemma 4. Note that

$$\begin{vmatrix} 1 & \phi(x_1) & \psi(x_1) \\ 1 & \phi(x_2) & \psi(x_2) \\ 1 & \phi(x_3) & \psi(x_3) \end{vmatrix} = (\phi(x_3) - \phi(x_2))(\phi(x_2) - \phi(x_1)) \left(\frac{\psi(x_3) - \psi(x_2)}{\phi(x_3) - \phi(x_2)} - \frac{\psi(x_2) - \psi(x_1)}{\phi(x_2) - \phi(x_1)} \right).$$

Since $\phi(x)$ is increasing in x, $\phi(x_1) \leq \phi(x_2) \leq \phi(x_3)$ for every $x_1 < x_2 < x_3$. Therefore, we have, for every $x_1 < x_2 < x_3$,

$$\begin{vmatrix} 1 & \phi(x_1) & \psi(x_1) \\ 1 & \phi(x_2) & \psi(x_2) \\ 1 & \phi(x_3) & \psi(x_3) \end{vmatrix} \underset{(\leq)}{\geq} 0 \quad \Longleftrightarrow \quad \frac{\psi(x_3) - \psi(x_2)}{\phi(x_3) - \phi(x_2)} - \frac{\psi(x_2) - \psi(x_1)}{\phi(x_2) - \phi(x_1)} \underset{(\leq)}{\geq} 0,$$

which indicates that $\psi(x)$ is convex (concave) in $\phi(x)$.

■

Proof of Lemma 5. Assume that $x_1 < x_2 < x_3$ and $a_1 < a_2 < a_3$, and, without lost of generality, let $a^o = a_2$.

(i) the "if" part: Since

$$T(f,2) \equiv \begin{vmatrix} f(x_1|a_1) & f(x_1|a_2) \\ f(x_2|a_1) & f(x_2|a_2) \end{vmatrix} = f(x_1|a_1)f(x_2|a_1) \times \begin{vmatrix} 1 & \dfrac{f(x_1|a_2)}{f(x_1|a_1)} \\ 1 & \dfrac{f(x_2|a_2)}{f(x_2|a_1)} \end{vmatrix}$$
$$= f(x_1|a_1)f(x_2|a_1)\left(\frac{f(x_2|a_2)}{f(x_2|a_1)} - \frac{f(x_1|a_2)}{f(x_1|a_1)}\right), \tag{A.34}$$

the MLRP for $f(x|a)$ implies $T(f,2) \geq 0$. Also, given the MLRP for $f(x|a)$, condition (ii) in Lemma 5 implies that $\frac{f(x|a_1)}{f(x|a^o)}$ is convex in $q = \frac{f_a(x|a^o)}{f(x|a^o)}$, $\forall a_1 < a^o$ as well as that $\frac{f(x|a_3)}{f(x|a^o)}$ is convex in $q = \frac{f_a(x|a^o)}{f(x|a^o)}$, $\forall a_3 > a^o$. Therefore, we have[18]

$$\frac{\frac{f(x_2|a_1)}{f(x_2|a^o)} - \frac{f(x_1|a_1)}{f(x_1|a^o)}}{q_2 - q_1} \leq \frac{\frac{f(x_3|a_1)}{f(x_3|a^o)} - \frac{f(x_2|a_1)}{f(x_2|a^o)}}{q_3 - q_2} \leq 0 \tag{A.35}$$

and

$$0 \leq \frac{\frac{f(x_2|a_3)}{f(x_2|a^o)} - \frac{f(x_1|a_3)}{f(x_1|a^o)}}{q_2 - q_1} \leq \frac{\frac{f(x_3|a_3)}{f(x_3|a^o)} - \frac{f(x_2|a_3)}{f(x_2|a^o)}}{q_3 - q_2}, \tag{A.36}$$

where $q_i = \frac{f_a(x_i|a^o)}{f(x_i|a^o)}$. By combining (A.35) and (A.36), we derive

$$\frac{\frac{f(x_2|a_1)}{f(x_2|a^o)} - \frac{f(x_1|a_1)}{f(x_1|a^o)}}{\frac{f(x_2|a_3)}{f(x_2|a^o)} - \frac{f(x_1|a_3)}{f(x_1|a^o)}} \leq \frac{\frac{f(x_3|a_1)}{f(x_3|a^o)} - \frac{f(x_2|a_1)}{f(x_2|a^o)}}{\frac{f(x_3|a_3)}{f(x_3|a^o)} - \frac{f(x_2|a_3)}{f(x_2|a^o)}} \leq 0. \tag{A.37}$$

By definition we can rewrite $T(f,3)$ as

$$T(f,3) \equiv \begin{vmatrix} f(x_1|a_1) & f(x_1|a^o) & f(x_1|a_3) \\ f(x_2|a_1) & f(x_2|a^o) & f(x_2|a_3) \\ f(x_3|a_1) & f(x_3|a^o) & f(x_3|a_3) \end{vmatrix} = \left(\prod_{i=1}^{3} f(x_i|a^o)\right) \times \begin{vmatrix} \frac{f(x_1|a_1)}{f(x_1|a^o)} & 1 & \frac{f(x_1|a_3)}{f(x_1|a^o)} \\ \frac{f(x_2|a_1)}{f(x_2|a^o)} & 1 & \frac{f(x_2|a_3)}{f(x_2|a^o)} \\ \frac{f(x_3|a_1)}{f(x_3|a^o)} & 1 & \frac{f(x_3|a_3)}{f(x_3|a^o)} \end{vmatrix}$$
$$= \left(\prod_{i=1}^{3} f(x_i|a^o)\right) \times \begin{vmatrix} 1 & \frac{f(x_1|a_3)}{f(x_1|a^o)} & \frac{f(x_1|a_1)}{f(x_1|a^o)} \\ 1 & \frac{f(x_2|a_3)}{f(x_2|a^o)} & \frac{f(x_2|a_1)}{f(x_2|a^o)} \\ 1 & \frac{f(x_3|a_3)}{f(x_3|a^o)} & \frac{f(x_3|a_1)}{f(x_3|a^o)} \end{vmatrix},$$

one can easily check that (A.37) implies $T(f,3) \geq 0$.

[18]Due to the MLRP, for $a_1 < a^o$, $\frac{f(x|a_1)}{f(x|a^o)}$ becomes decreasing in $\forall x$.

(ii) the "only if" part: If $f(x|a)$ is TP_3, then, by definition, $T(f,2) \geq 0$ and $T(f,3) \geq 0$. First, from (A.34), it is obvious that $T(f,2) \geq 0$ implies the MLRP for $f(x|a)$. Second, notice that

$$\begin{aligned}
T(f,3) &\equiv \begin{vmatrix} f(x_1|a_1) & f(x_1|a^o) & f(x_1|a_3) \\ f(x_2|a_1) & f(x_2|a^o) & f(x_2|a_3) \\ f(x_3|a_1) & f(x_3|a^o) & f(x_3|a_3) \end{vmatrix} \\
&= (a^o - a_1) \times \begin{vmatrix} f(x_1|a_1) & \frac{f(x_1|a^o)-f(x_1|a_1)}{a^o-a_1} & f(x_1|a_3) \\ f(x_2|a_1) & \frac{f(x_2|a^o)-f(x_2|a_1)}{a^o-a_1} & f(x_2|a_3) \\ f(x_3|a_1) & \frac{f(x_3|a^o)-f(x_3|a_1)}{a^o-a_1} & f(x_3|a_3) \end{vmatrix} \\
&= (a^o - a_1) \times \left\{ \prod_{i=1}^{3} f(x_i|a_1) \right\} \times \underbrace{\begin{vmatrix} 1 & \frac{f(x_1|a^o)-f(x_1|a_1)}{(a^o-a_1)f(x_1|a_1)} & \frac{f(x_1|a_3)}{f(x_1|a_1)} \\ 1 & \frac{f(x_2|a^o)-f(x_2|a_1)}{(a^o-a_1)f(x_2|a_1)} & \frac{f(x_2|a_3)}{f(x_2|a_1)} \\ 1 & \frac{f(x_3|a^o)-f(x_3|a_1)}{(a^o-a_1)f(x_3|a_1)} & \frac{f(x_3|a_3)}{f(x_3|a_1)} \end{vmatrix}}_{\equiv A}.
\end{aligned}$$

Since $a^o > a_1$ and since

$$\lim_{a_1 \to a^o} A = \begin{vmatrix} 1 & \frac{f_a(x_1|a^o)}{f(x_1|a^o)} & \frac{f(x_1|a_3)}{f(x_1|a^o)} \\ 1 & \frac{f_a(x_2|a^o)}{f(x_2|a^o)} & \frac{f(x_2|a_3)}{f(x_2|a^o)} \\ 1 & \frac{f_a(x_3|a^o)}{f(x_3|a^o)} & \frac{f(x_3|a_3)}{f(x_3|a^o)} \end{vmatrix},$$

$T(f,3) \geq 0$ implies that $\frac{f(x|a_3)}{f(x|a^o)}$ is convex in $q = \frac{f_a(x|a^o)}{f(x|a^o)}$, $\forall a_3 > a^o$ by Lemma 4.

Similarly, notice that

$$\begin{aligned}
T(f,3) &= (a_3 - a^o) \times \begin{vmatrix} f(x_1|a_1) & f(x_1|a^o) & \frac{f(x_1|a_3)-f(x_1|a^o)}{a_3-a^o} \\ f(x_2|a_1) & f(x_2|a^o) & \frac{f(x_2|a_3)-f(x_2|a^o)}{a_3-a^o} \\ f(x_3|a_1) & f(x_3|a^o) & \frac{f(x_3|a_3)-f(x_3|a^o)}{a_3-a^o} \end{vmatrix} \\
&= (a_3 - a^o) \times \left(\prod_{i=1}^{3} f(x_i|a^o) \right) \times \begin{vmatrix} \frac{f(x_1|a_1)}{f(x_1|a^o)} & 1 & \frac{f(x_1|a_3)-f(x_1|a^o)}{(a_3-a^o)f(x_1|a^o)} \\ \frac{f(x_2|a_1)}{f(x_2|a^o)} & 1 & \frac{f(x_2|a_3)-f(x_2|a^o)}{(a_3-a^o)f(x_2|a^o)} \\ \frac{f(x_3|a_1)}{f(x_3|a^o)} & 1 & \frac{f(x_3|a_3)-f(x_3|a^o)}{(a_3-a^o)f(x_3|a^o)} \end{vmatrix} \\
&= (a_3 - a^o) \times \left(\prod_{i=1}^{3} f(x_i|a^o) \right) \times \underbrace{\begin{vmatrix} 1 & \frac{f(x_1|a_3)-f(x_1|a^o)}{(a_3-a^o)f(x_1|a^o)} & \frac{f(x_1|a_1)}{f(x_1|a^o)} \\ 1 & \frac{f(x_2|a_3)-f(x_2|a^o)}{(a_3-a^o)f(x_2|a^o)} & \frac{f(x_2|a_1)}{f(x_2|a^o)} \\ 1 & \frac{f(x_3|a_3)-f(x_3|a^o)}{(a_3-a^o)f(x_3|a^o)} & \frac{f(x_3|a_1)}{f(x_3|a^o)} \end{vmatrix}}_{\equiv B}.
\end{aligned}$$

Since $a_3 > a^o$ and since

$$\lim_{a_3 \to a^o} B = \begin{vmatrix} 1 & \frac{f_a(x_1|a^o)}{f(x_1|a^o)} & \frac{f(x_1|a_1)}{f(x_1|a^o)} \\ 1 & \frac{f_a(x_2|a^o)}{f(x_2|a^o)} & \frac{f(x_2|a_1)}{f(x_2|a^o)} \\ 1 & \frac{f_a(x_3|a^o)}{f(x_3|a^o)} & \frac{f(x_3|a_1)}{f(x_3|a^o)} \end{vmatrix},$$

$T(f,3) \geq 0$ also implies that $\frac{f(x|a_1)}{f(x|a^o)}$ is convex in $q = \frac{f_a(x|a^o)}{f(x|a^o)}$, $\forall a_1 < a^o$ by Lemma 4. Consequently, $T(f,3) \geq 0$ implies that, for any given a^o, $\frac{f(x|a)}{f(x|a^o)}$ is convex in $q = \frac{f_a(x|a^o)}{f(x|a^o)}$ for all a.

■

Proof of Corollary 6:. We have, for any given a^o,

$$\begin{aligned} T(f,2) \equiv \begin{vmatrix} f(x_1|a_1) & f(x_1|a_2) \\ f(x_2|a_1) & f(x_2|a_2) \end{vmatrix} &= Q'_{a^o}(x_1)Q'_{a^o}(x_2) \begin{vmatrix} g(q_1|a_1) & g(q_1|a_2) \\ g(q_2|a_1) & g(q_2|a_2) \end{vmatrix} \\ &= Q'_{a^o}(x_1)Q'_{a^o}(x_2)T(g,2), \end{aligned}$$

where $q_i = Q_{a^o}(x_i) \equiv \frac{f_a(x_i|a^o)}{f(x_i|a^o)}$, $i = 1,2$. Since $f(x|a)$ satisfies MLRP, we have $Q'_{a^o}(x) \geq 0, \forall x$. Thus, given MLRP for $f(x|a)$,

$$\begin{aligned} & T(f,2) \geq 0, \quad \text{for every } x_1 < x_2 \text{ and } a_1 < a_2 \\ \Longleftrightarrow\ & T(g,2) \geq 0, \quad \text{for every } q_1 < q_2 \text{ and } a_1 < a_2. \end{aligned} \tag{A.38}$$

Likewise, we have, for any given a^o,

$$\begin{aligned} T(f,3) &\equiv \begin{vmatrix} f(x_1|a_1) & f(x_1|a_2) & f(x_1|a_3) \\ f(x_2|a_1) & f(x_2|a_2) & f(x_2|a_3) \\ f(x_3|a_1) & f(x_3|a_2) & f(x_3|a_3) \end{vmatrix} \\ &= Q'_{a^o}(x_1)Q'_{a^o}(x_2)Q'_{a^o}(x_3) \begin{vmatrix} g(q_1|a_1) & g(q_1|a_2) & g(q_1|a_3) \\ g(q_2|a_1) & g(q_2|a_2) & g(q_2|a_3) \\ g(q_3|a_1) & g(q_3|a_2) & g(q_3|a_3) \end{vmatrix} \\ &= Q'_{a^o}(x_1)Q'_{a^o}(x_2)Q'_{a^o}(x_3)T(g,3), \end{aligned}$$

where $q_i = Q_{a^o}(x_i) \equiv \frac{f_a(x_i|a^o)}{f(x_i|a^o)}$, $i = 1,2,3$. Therefore, by the same way, we derive that, given MLRP for $f(x|a)$,

$$\begin{aligned} & T(f,3) \geq 0, \quad \text{for every } x_1 < x_2 < x_3 \text{ and } a_1 < a_2 < a_3 \\ \iff \quad & T(g,3) \geq 0, \quad \text{for every } q_1 < q_2 < q_3 \text{ and } a_1 < a_2 < a_3 \end{aligned} \tag{A.39}$$

Thus, by combining (A.38) and (A.39), we have that $f(x|a)$ is TP_3 if and only if $g(q|a)$ is TP_3 for any given a^o given that $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP.

■

Proof of Proposition 5. From Lemma 5 and Corollary 1, it is shown that, given that $x \in \mathbb{R}$ and $f(x|a)$ satisfies MLRP, condition (TP_3) is equivalent to condition (1a). Furthermore, condition (1J-2) and condition (2b) are equivalent.

■

Proof of Lemma 7. Define, for a given density function $f(x|a)$,

$$\psi(a,k) \equiv \int \phi(x,k) f(x|a) dx,$$

where $x \in \mathbb{R}$, $a \in \mathbb{R}$, and k is the parameter that determines the functional form of $\phi(x,k)$. Then, by the "basic composition formula" by Karlin (1968),[19] we have

$$\begin{vmatrix} \psi(a_1,k_1) & \psi(a_1,k_2) & \psi(a_1,k_3) \\ \psi(a_2,k_1) & \psi(a_2,k_2) & \psi(a_2,k_3) \\ \psi(a_3,k_1) & \psi(a_3,k_2) & \psi(a_3,k_3) \end{vmatrix} = \iiint_{x_1<x_2<x_3} \begin{vmatrix} \phi(x_1,k_1) & \phi(x_1,k_2) & \phi(x_1,k_3) \\ \phi(x_2,k_1) & \phi(x_2,k_2) & \phi(x_2,k_3) \\ \phi(x_3,k_1) & \phi(x_3,k_2) & \phi(x_3,k_3) \end{vmatrix} \times \begin{vmatrix} f(x_1|a_1) & f(x_1|a_2) & f(x_1|a_3) \\ f(x_2|a_1) & f(x_2|a_2) & f(x_2|a_3) \\ f(x_3|a_1) & f(x_3|a_2) & f(x_3|a_3) \end{vmatrix} dx_1 dx_2 dx_3. \tag{A.40}$$

Let $\phi(x,k_1) \equiv 1$, $\phi(x,k_2) \equiv x$, and $\phi(x,k_3) \equiv u(x)$. We have $\psi(a,k_1) = 1$, $\psi(a,k_2) = \mu(a) \equiv \int x f(x|a) dx$, and $\psi(a,k_3) = u^*(a) \equiv \int u(x) f(x|a) dx$. Thus, by using (A.40),

$$\begin{vmatrix} 1 & \mu(a_1) & u^*(a_1) \\ 1 & \mu(a_2) & u^*(a_2) \\ 1 & \mu(a_3) & u^*(a_3) \end{vmatrix} = \iiint_{x_1<x_2<x_3} \begin{vmatrix} 1 & x_1 & u(x_1) \\ 1 & x_2 & u(x_2) \\ 1 & x_3 & u(x_3) \end{vmatrix} \times \begin{vmatrix} f(x_1|a_1) & f(x_1|a_2) & f(x_1|a_3) \\ f(x_2|a_1) & f(x_2|a_2) & f(x_2|a_3) \\ f(x_3|a_1) & f(x_3|a_2) & f(x_3|a_3) \end{vmatrix} dx_1 dx_2 dx_3. \tag{A.41}$$

[19]For the detailed proof of the basic composition formula, see Karlin (1968) p. 17. The (A.40) is a direct extension of the famous Cauchy-Binet theorem.

The fact that $f(x|a)$ is TP_3 implies

$$T(f,2) = \begin{vmatrix} f(x_1|a_1) & f(x_1|a_2) \\ f(x_2|a_1) & f(x_2|a_2) \end{vmatrix} \geq 0, \qquad \text{for every } x_1 < x_2 \text{ and } a_1 < a_2,$$

which is equivalent to the MLRP for $f(x|a)$. Thus, both $\mu(a)$ and $u^*(a)$ are increasing in a. Since $u(x)$ is concave in x, we have from Lemma 4 that

$$\begin{vmatrix} 1 & x_1 & u(x_1) \\ 1 & x_2 & u(x_2) \\ 1 & x_3 & u(x_3) \end{vmatrix} \leq 0, \qquad \text{for every } x_1 < x_2 < x_3.$$

Also, the fact that $f(x|a)$ is TP_3 implies

$$T(f,3) = \begin{vmatrix} f(x_1|a_1) & f(x_1|a_2) & f(x_1|a_3) \\ f(x_2|a_1) & f(x_2|a_2) & f(x_2|a_3) \\ f(x_3|a_1) & f(x_3|a_2) & f(x_3|a_3) \end{vmatrix} \geq 0, \qquad \text{for every } x_1 < x_2 < x_3 \text{ and } a_1 < a_2 < a_3.$$

Thus, from (A.41), we have

$$\begin{vmatrix} 1 & \mu(a_1) & u^*(a_1) \\ 1 & \mu(a_2) & u^*(a_2) \\ 1 & \mu(a_3) & u^*(a_3) \end{vmatrix} \leq 0, \qquad \text{for every } a_1 < a_2 < a_3,$$

which indicates that $u^*(a)$ is increasing concave in $\mu(a) \equiv \int x f(x|a)dx$, given that $\mu(a)$ is increasing in a.

■

B.4 Additional Results

Derivation of (22) in Example 3. Given

$$g(q|a) = \frac{[h(a^o)]^2}{h'(a^o)h(a)} \exp\left(-\frac{1}{h(a)}\left(\frac{[h(a^o)]^2}{h'(a^o)}q + h(a^o)\right)\right),$$

the moment generating function for arbitrary a will be given as

$$\begin{aligned}
M(a;t) &= \int e^{tq} g(q|a) dq = \frac{[h(a^o)]^2}{h'(a^o)h(a)} \int_{-\frac{h'(a^o)}{h(a^o)}}^{\infty} \exp\left[-\frac{1}{h(a)}\left(\frac{h(a^o)^2}{h'(a^o)} q + h(a^o)\right) + tq\right] dq \\
&= \frac{[h(a^o)]^2}{h'(a^o)h(a)} \exp\left(-\frac{h(a^o)}{h(a)}\right) \int_{-\frac{h'(a^o)}{h(a^o)}}^{\infty} \exp\left[-\left(\frac{h(a^o)^2}{h(a)h'(a^o)} - t\right) q\right] dq \\
&= \frac{[h(a^o)]^2}{h'(a^o)h(a)} \frac{1}{\dfrac{h(a^o)^2}{h(a)h'(a^o)} - t} \exp\left(-t\frac{h'(a^o)}{h(a^o)}\right) = \frac{h(a^o)^2}{h(a^o)^2 - h'(a^o)h(a)t} \exp\left(-t\frac{h'(a^o)}{h(a^o)}\right),
\end{aligned}$$

which is (22). ■