Agency Theory and Higher Order Stochastic Dominance

Claudio A. Bonilla, University of Chile*
Hervé Roche, Universidad Adolfo Ibáñez
Marcos Vergara, Universidad del Desarrollo

December 15, 2020

Abstract

We study the classic agency model when the suitable ordering of risky prospects is better represented by higher rather than first-order stochastic dominance. We derive conditions on preferences under which if the effect of the agent's effort corresponds to a Nth-order stochastic dominance improvement, the principal would want a greater effort level provided by the agent than the one provided under the second best sharing rule. Our characterization involves the inverses of the marginal utilities and the ratios of their successive derivatives up to the $N-1$ order. Finally, we illustrate our findings for standard preferences used in financial economics.

Keywords: Stochastic dominant shifts, agency theory, higher-order risk attitudes.
JEL Classification: D81, D86, M2.

*Corresponding author, cbonilla@fen.uchile.cl. Diagonal Paraguay 257, suite 1905, Santiago, Chile.
We thank the comments of seminar participants at the University of Chile, Universidad Adolfo Ibáñez, and at the 13th International Conference of the Western Economic Association Meeting in Santiago de Chile in 2017.

1 Introduction

Agency models have become a prominent tool to analyze a wide range of economic situations. These models were initially developed to study insurance markets (Borch, 1962; Zeckhauser, 1970), in the economic analyses of incentives and authority within organizations (Mirrlees, 1976), and also in the study of financial markets (Jensen and Meckling, 1976; Myers, 1977 and Ross, 1977). But, as the academic community observed the usefulness of the approach, other sub-fields of economics soon began to incorporate agency models into their analyses. Today, from accounting to labor relations, and from corporate finance to macroeconomics, most economic specialties have included the principal-agent perspective as part of their theoretical framework of analysis.

One key characteristic of agency problems is that the principal (the individual who offer a contract) would always wants the agent (the individual who executes a task) to exert more effort than the actual effort exerted in the second-best equilibrium (Hölmstrom, 1979). This is observed through a positive shadow price of the agent's participation constraint. The economic intuition of this result is based on the conditional probability distribution of the output, which it is defined in a way that guarantee that "the greater the effort exerted by the agent, the **better** the distribution of results turns out". In consequence, defining the changes in the conditional distribution of results when the agent's effort changes is a key element in the modeling agency problems.

In this paper we investigate how higher-order stochastic changes in the risky distribution of results due to changes in the levels of agent's effort, can be incorporated into traditional agency problems. More specifically, we derive sufficient conditions that involve the inverse functions of the marginal utility functions of the principal and the agent under which the principal would like to see the agent to exert a higher level of effort than the one given by the second best sharing rules. This method allows us to capture the mathematical properties of these sufficient conditions and, at the same time, go beyond the traditional (and limiting) case of first-order stochastic dominance (FOSD) shifts in the distribution of results. We study the optimal contract when increases in agent's efforts translate not only into the usual FOSD shift, but also into stochastic changes of higher orders of the conditional distribution of results. We exemplify our results and tests their robustness using traditional utility functions from the financial economics literature.

In the last twenty years we have seen important developments in the economic theory of risk, and in particular in the theory of higher-order risk analysis. Therefore, we believe that incorporating these features into traditional applied economic models that have basic risk structures, like the agency problem, is one way of enriching those models and to gain important new insights of traditional applied economic problems.

The standard way to think about the effects of agent's effort on the distribution of results is to consider a FOSD shifts. Under FOSD shifts, as long as the utility function of an economic agent is increasing in wealth, if distribution A is always preferred over distribution B, then distribution A fist-order stochastic dominates distribution B (see Mas-Colell et al., 1995, page 195). The literature shows that FOSD is a universal assumption in agency problems, and this is the case because FOSD shifts is a natural and intuitive way to think about better distributions of results.

But FOSD shifts are not the only way to represent better distributions of results. When economic agents are highly risk averse and volatility is a major issue, second-order stochastic dominance (SOSD) shifts may be more important for the principal in some specific situations. For example, let us suppose that the principal is delegating to an agent the management of a portfolio of new ventures. If the principal is about to retire and therefore he is thinking in selling his portfolio of ventures in near future, he may want his investments to be managed in a way that minimizes risk (or volatility) for a given level of expected returns, Then, instead of the usual assumption of maximizing expected values (for risk neutral principals) in agency problems, the principal may want that the increase in the agent's effort be translated into safer investments, that is, investing in new ventures that decrease total risk of the portfolio of firms or alternatively, that the additional effort implies increasing control over the key performance indicators of the venture in order to minimize the volatility of the venture's results. This approach can certainly be described by a SOSD shift instead of a FOSD shift. By way of illustration, a special case of SOSD shift is a mean-preserving decrease in risk described by Rothschild and Stiglitz (1970). Therefore, in the context just described, the intuitive way to model increases in agent's effort is producing SOSD shifts but this feature is not typically included in agency problems.

But we have yet more cases. Consider the case where the principal's primary objective is to avoid bad states of nature at all cost. If this were the case, the principal can be characterized as an individual who exhibits high downside risk aversion (Menezes et. al, 1980). Downside risk aversion implies that the principal is prudent (Kimball, 1990) and therefore, if we assume that at different effort levels the first two statistical moments of the distribution of results are equal, then third-order stochastic dominant (TOSD) shifts will provide the right ordering of risky prospects. Another example is when the principal is temperate (Kimball [1993]) and thus he exhibits aversion to outer risk (Menezes and Wang [2005]) or equivalently the principal dislike kurtosis. In this case, fourth-order stochastic dominant (FOOSD) shifts will provide the right ordering of distributions. We can continue for higher-order risk changes infinitely. However, the point we are making is that, by incorporating higher-order risk structures, we can go beyond the traditional

FOSD perspective of agency problems, increasing the spectrum of economic situations that can be studied and, the same time, providing more flexibility for modeling new types of agency games.

In terms of the literature we are advancing, this paper connects traditional the agency model with recent developments in the higher–order risk literature. In the higher-order risk literature, our paper is related with several papers that also advance applied economic problems with higher-order risk structures. For instance, Jindapon and Nielsen (2007) develop a higher-order generalization of Arrow-Pratt and Ross (1981) risk aversion coefficients. Jouini et al. (2013) study comparative statistics of Nth-degree risk increases in bi-dimensional problems with additive and multiplicative risks and provide the conditions for unambiguous impact on optimal decisions. Liu and Meyer (2013) study higher-order risk through defining a rate of substitution of one stochastic change to a random variable to another. A very good summary containing recent developments of risk aversion measures under higher-order risk is given in Liu and Nielsen (2019).

In the agency literature, our paper is related with most of the papers that build from the first-order approach (FOA) of Holmström (1979). For instance, Jewitt (1988), provided the conditions that make valid the FOA in the multi-signal case avoiding the convexity of the distribution function (CDF) condition. Sinclair-Desgasné (1994) provided a generalization of the CDF, that is called the general convexity of the distribution function (GCDF) for the multi-signal case. Alvi (1997) showed that the FOA is valid only for additively separable utility functions and provided the sufficient conditions to extend the FOA to non-separable cases. Licalzi and Speater (2003) provided a family of examples that satisfy the monotone likelihood ration condition and the CDF condition.

More recently, other theoretical approaches in agency theory have worked on improving different aspects of the principal-agent model. For instance, Conlon (2009) presents a multi-signal generalization of the methods developed in Rogerson (1985) and Jewitt (1988), and he compares the cases in which one method works better than the other. Sannikov (2008) presents a continuous-time version of the principal-agent problem in which the output follows a diffusion process. Cato (2013) identifies the conditions under which the FOA is valid under inequality aversion and Abraham et al. (2011) provide sufficient conditions for the validity of the FOA for a two-period dynamic moral hazard model with hidden borrowing and lending. Finally, Chaigneau et al (2019) extend the Holmstrom (1979) paper by providing new conditions for a signal to have a positive value when the FOA does not hold.

The rest of the paper is organized as follows. Section 2 describes the economic framework of the standard agency problem. In section 3 we discuss stochastic dominance shifts induced by higher effort's levels in the agency context. In this section we also characterize preferences that satisfy the mixed risk aversion property. Section 4 studies the effects of first, second and Nth-order stochastic dominance shifts and look for conditions on preferences the guarantee that the principal wishes the agent to increase his effort. In section 5, we illustrate our results with standard utility functions used in the financial economics literature. Finally, section 6 provides the conclusions. All proofs are contained in the appendix.

2 The Model

We study a principal-agent problem where the agent privately takes an action a from set $A \subseteq \mathbb{R}$. The monetary payoff $\widetilde{x}$ is random and let $x \in X \subseteq \mathbb{R}$ denote an outcome. Without loss of generality, we assume that X is an interval of the form $[\underline{x}, \overline{x}]$, with $\underline{x} < \overline{x}$ and $(\underline{x}, \overline{x}) \in \overline{\mathbb{R}} \times \overline{\mathbb{R}}$. The agent first decides his action a (we can think of a as an effort level) and then both the principal and the agent observe outcome x. We denote by $F(\cdot \mid a)$ the cumulative distribution function of $\widetilde{x}$, given action a taken by the agent. In addition, we assume that the support of F is independent of a and there exists a conditional density function denoted by $f(\cdot \mid a)$ such that $f_a(x \mid a)$ and $f_{aa}(x \mid a)$ represent the first and second order derivative of $f(\cdot \mid a)$ and they are well-defined for all (x, a). Finally, let $s(x)$ be the share of the payoff received by the agent (from and optimal sharing rule that we will present in next section) given payoff x, and let $r(x) = x - s(x)$ be the share of the payoff received by the principal. Following Hölmstrom (1979), given outcome x, we restrict $s(x)$ to lie in some interval I of the form $[c, x + d]$ in order to avoid non-existence of a solution.

2.1 The Agent

The agent's preferences are represented by a separable expected utility function that is given by $\int u(s(x))f(x \mid a)dx - c(a)$, where u is a Von Neumann-Morgenstern utility function defined on set Δ_u with $u' > 0$, $u'' < 0$, $u''' > 0$, $u'''' < 0$. That is, the agent exhibits positive marginal utility, risk aversion, prudence (aversion to downside risk) and temperance (aversion to outer risk) and $c(a)$ is a convex cost function that represents the disutility of the agent's effort and we assume that $c' > 0$ and $c'' > 0$. In addition, the agent has an outside job opportunity that provides reservation utility u_0.

Given a sharing rule $s(x)$, the agent aims to maximize his expected utility

$$\max_{a \in A} \int_X u(s(x))f(x \mid a)dx - c(a).$$

Assuming an interior solution, the optimal effort a^0 that the agent exerts must be such that

$$\int_X u(s(x))f_a(x \mid a^0)dx - c'(a^0) = 0,$$

which represents his optimal choice of a given the sharing rule $s(x)$. We assume that our generic functional forms satisfy the monotone likelihood ratio property (MLR property) and the convexity of the distribution function. Therefore, the first order approach (FOA) can be used in our problem in order to find the optimal contract [1].

2.2 The Principal

Given an action a taken by the agent, the principal's expected utility function is given by

$$\int_X G\left(x - s\left(x\right)\right) f\left(x \mid a\right) dx,$$

where G is a also a Von Neumann-Morgenstern utility function defined on set Δ_G with $G' > 0$, $G'' < 0$, $G''' > 0$ and $G^{iv} < 0$. That is, the principal also exhibits positive marginal utility, risk aversion, prudence and temperance, and since the principal's sharing rule r was defined as $r(x) = x - s(x)$, the principal's optimal decision problem is:

$$\begin{array}{lll} \max\limits_{\{s \in I,\ a \in A\}} & \int_X G\left(r\left(x\right)\right) f\left(x \mid a\right) dx & \\ \text{s.t.} & \int_X u(s(x)) f\left(x \mid a\right) dx - c\left(a\right) \geq u_0 & \text{(C1)} \\ & \int_X u(s(x)) f_a\left(x \mid a\right) dx - c'\left(a\right) = 0 & \text{(C2)} \end{array}$$

This problem can be solved using the standard Lagrange method. Let λ and μ denote the Lagrange multipliers of constraints (C1) and (C2) respectively. Constraint (C1) is the participation constraint (or the individual rationality constraint) whereas constraint (C2) is the agent's incentive compatibility constraint.

[1] Agency models have always presented some technical difficulties. One of the main concerns of these models has been about the cases and situations in which it is correct to use the first order approach (FOA). Mirrlees (1975) was the first to point out that the FOA is generally not valid and that more assumptions – in particular two additional assumptions - are needed in order to achieve a correct solution when using the FOA. These assumptions (or required conditions) are described in detail in Mirrlees (1975) and Rogerson (1985) and they are the monotone likelihood ratio (MLR) property and the convexity of the distribution function (CDF). The first condition is very intuitive and it can be interpreted as "more effort induces better results", in line with the intuition of FOSD shifts. The second condition is more technical and it is required by the mathematical optimization program used in agency modeling in order to achieve an optimal solution

2.3 Optimality Conditions

2.3.1 Pareto Rule

When there is no asymmetric information, the principal observes a and can contract directly based on a. In this case, we do not need the incentive compatibility constraint. Then, using pointwise optimization, the first order condition of the Lagrangian function L with respect to $s(x)$ is given by:

$$\frac{\partial L}{\partial s} = -G'\left(r\left(x\right)\right) f\left(x \mid a\right) + \lambda u'\left(s\left(x\right)\right) f\left(x \mid a\right) = 0, \tag{1}$$

so the optimality condition of equation (1) becomes

$$\frac{G'\left(r^*_{\lambda^*}\left(x\right)\right)}{u'\left(s^*_{\lambda^*}\left(x\right)\right)} = \lambda^* > 0. \tag{2}$$

This optimal sharing rule $s^*_{\lambda^*}$ is called the *Borch rule* and we will refer to it as the *Pareto optimal solution* (*first best* case). Furthermore, the first order condition for a is given by:

$$\int_X \left[G\left(r^*_{\lambda^*}\left(x\right)\right) + \lambda^* u\left(s^*_{\lambda^*}\left(x\right)\right)\right] f_a\left(x \mid a\right) dx = \lambda^* c'(a).$$

Normalization of λ^*. It is possible to normalize to 1 the Lagrange multiplier λ^*. Since under expected utility preferences can be represented by a utility function up to a positive affine transformation, we may consider for the agent utility function $u_\theta(x) = u(x) - \theta$, with $\theta \in \mathbb{R}$. From relationship (2), we can obtain a formal expression of s^* as a function of (x, λ^*), i.e. $s^*(x, \lambda^*)$. Using the implicit function theorem we obtain that

$$\frac{\partial s^*(x,\lambda^*)}{\partial \lambda^*} = -\frac{u'\left(s^*\left(x,\lambda^*\right)\right)}{u''\left(s^*\left(x,\lambda^*\right)\right) + \dfrac{G'(x - s^*\left(x,\lambda^*\right))}{u'(s^*\left(x,\lambda^*\right))} G''\left(x - s^*\left(x,\lambda^*\right)\right)} > 0,$$

so we can claim that s^* is a function of parameter λ^* that is increasing. Then plugging $s^*(x, \lambda^*)$ into relationship C1 leads to

$$\begin{aligned} \int_X u_\theta\left(s^*\left(x,\lambda^*\right)\right) f\left(x \mid a\right) dx - c\left(a\right) &= u_0 \\ \int_X u\left(s^*\left(x,\lambda^*\right)\right) f\left(x \mid a\right) dx &= u_0 + c\left(a\right) + \theta. \end{aligned}$$

The left handside the previous equality is increasing in λ^* and the right handside of the equality as a function of θ can take any value. Choosing parameter θ such that

$$\int_X u\left(s^*\left(x,1\right)\right) f\left(x \mid a\right) dx = u_0 + c\left(a\right) + \theta$$

will ensure that indeed $\lambda^* = 1$.

In the remainder of the paper, we normalize[2] the value of Lagrange multiplier λ^* to be equal to 1 and denote by r^* and s^* the Pareto optimal sharing rules for the normalized problem.

Finally, normalizing $\lambda^* = 1$, from relationship (2) it follows that $x - r^*(x) = I_u(G'(r^*(x))$ where I_u is the inverse function of u'. Thus, we have

$$\Gamma_u(r^*(x)) = x,$$

where function Γ_u is defined as

$$\Gamma_u(y) = y + I_u(G'(y)). \tag{3}$$

Γ_u is a smooth function with $\Gamma_u'(y) = 1 + \dfrac{G''(y)}{u''(I_u(G'(y)))} > 0$, so by the Implicit Function Theorem function Γ_u admits an inverse on X denoted I_{Γ_u}. Finally, we obtain that $r^*(x) = I_{\Gamma_u}(x)$, i.e., a relationship between the principal's optimal share of the payoff and the utility functions of the principal and the agent.

2.3.2 Moral Hazard Rule

Under asymmetric information, the incentive compatibility constraint now has to be taken in account. Using pointwise optimization, the first order condition of the Lagrangian function L with respect to s is given by:

$$\frac{\partial L}{\partial s} = -G'\left(r\left(x\right)\right) f\left(x \mid a\right) + \lambda u'\left(s\left(x\right)\right) f\left(x \mid a\right) + \mu\left(u'\left(s\left(x\right)\right) f_a\left(x \mid a\right)\right) = 0,$$

which leads to

$$\frac{G'\left(r^{**}\left(x\right)\right)}{u'\left(s^{**}\left(x\right)\right)} = \lambda + \mu \frac{f_a\left(x \mid a\right)}{f\left(x \mid a\right)}. \tag{4}$$

s^{**} is the optimal sharing rule under asymmetric information (or under *moral hazard)* we will refer to it as the *second best* case.

Then, the first order condition of the Lagrangian with respect to a is given by

$$\begin{aligned}\frac{\partial L}{\partial a} = \int_X G\left(r\left(x\right)\right) f_a\left(x \mid a\right) dx + \lambda\left[\int_X u\left(s\left(x\right)\right) f_a\left(x \mid a\right) dx - c'\left(a\right)\right] \\ +\mu\left[\int_X u\left(s\left(x\right)\right) f_{aa}\left(x \mid a\right) dx - c''\left(a\right)\right] = 0,\end{aligned}$$

which leads to

$$\mu = -\frac{\int_X G\left(r^{**}\left(x\right)\right) f_a\left(x \mid a\right) dx}{\int_X u\left(s^{**}\left(x\right)\right) f_{aa}\left(x \mid a\right) dx - c''\left(a\right)}. \tag{5}$$

[2]Of course, the normalization shall affect the expressions of the optimal sharing rules but not the properties of their derivatives, which is the main focus of the paper.

Notice that the denominator of the expression (5), i.e., $\int_X u(s^{**}(x)) f_{aa}(x \mid a) dx - c''(a)$, is the second order condition of the agent's optimal effort and therefore, it is always negative given the assumptions of the model. If $\mu \geq 0$, it means that the principal wants the agent to exert more effort than the actual effort exerted in the second best equilibrium. But as we have argued, more effort can translate not only into FOSD shifts in the distribution of results, but also into higher order stochastic dominant shifts. For now, observe that in order to have $\mu \geq 0$, we only need $\int_X G(r^{**}(x)) f_a(x \mid a) dx$ to be positive. We shall get back to this point later in the paper.

Next, we present the concept of stochastic dominance of N^{th}-order, higher order risk aversion and mixed risk aversion to introduce the tools needed for our results.

3 Stochastic Dominance

3.1 Preliminary Concepts

We shall use the definitions given by Ekern (1980) to introduce the concepts of N^{th} order stochastic dominance and having less N^{th} degree risk. For $x \geq \underline{x}$, let us define $F^k(x \mid a) = \int_{\underline{x}}^x F^{k-1}(y \mid a) dy$, for $k = 1, 2, ..., N$. Note that $F^1(x \mid a) = F(x \mid a) = \int_{\underline{x}}^x F^0(y \mid a) dy$ and by definition $F^0(\cdot \mid a)$ is the density function $f(\cdot \mid a)$. Also, $F(\underline{x} \mid a) = 0$ and $F(\overline{x} \mid a) = 1$ for all $a \geq 0$. By iteration, we obtain

$$F^N(x \mid a) = \int_{\underline{x}}^x F^{N-1}(y \mid a) dy.$$

Definition 1. *Stochastic Dominance. Let F_1 and F_2 be two cumulative probability functions. We say the distribution F_1 N^{th} order stochastic dominates F_2 if we have:*

$$\begin{aligned} F_1^k(\overline{x}) &\leq F_2^k(\overline{x}), \text{ for } k = 1, 2, ..., N-1 \quad \text{and} \\ F_1^N(x) &\leq F_2^N(x), \text{ for } x \in X, \text{ with at least one strict inequality.} \end{aligned}$$

Definition 2. *Less N^{th} Degree Risk. Let F_1 and F_2 be two cumulative probability functions. We say the distribution F_1 has N^{th} less degree risk than F_2 if we have:*

$$\begin{aligned} F_1^N(x) &\leq F_2^N(x), \text{ for } x \in X, \text{ with at least one strict inequality, and} \\ F_1^k(\overline{x}) &= F_2^k(\overline{x}), \text{ for } k = 1, 2, ..., N \end{aligned}$$

This definition implies that the $N-1$ moments of these two distributions are the same and the difference between them is given by the N^{th} order distribution of the results. Therefore, we can say that having less N^{th} degree risk is a special case of N^{th} order stochastic dominance.

Definition 3. *Risk Aversion of N^{th} Order. Let u be a $C(\mathbb{R}_+)\cap C^N(\mathbb{R}_{++})$ utility function. An individual is said to exhibit N^{th} order risk aversion if*

$$(-1)^N u^{(N)}(x) < 0, \text{ for all } x \in \mathbb{R}_+. \tag{6}$$

Function u' is $(N+1)$-times monotone[3] on $\mathbb{R}_{++}$ and by Williamson's theorem (1956), u' admits the following representation

$$u'(x) = \int_0^\infty \left([1-xs]^+\right)^N dF_u(s),$$

for some distribution function F and where $[x]^+ = \max\{x, 0\}$.

3.2 Mixed Risk Aversion

In this section, we assume that the set of outcomes X is a subset of $\mathbb{R}_+$.

Definition 4. *Mixed Risk Aversion. An individual is said to exhibit "mixed risk aversion" if his preferences are represented by a $C^\infty(\mathbb{R}_+)$ utility function u such*

$$(-1)^n u^{(n)}(x) < 0, \text{ for all } n \in \mathbb{N} \text{ and } x \in \mathbb{R}_+$$

where $u^{(k)}$ denote the k^{th} derivative of u for $k \in \mathbb{N}$. Then, $A_n^u = -\frac{u^{(n+1)}}{u^{(n)}}$ is the coefficient of absolute risk aversion of order n, and it is well defined for al $n \in \mathbb{N}$.

Note that A_1^u is the Arrow-Pratt coefficient of absolute risk aversion, $A_2^u > 0$ is the prudence ratio that connects with precautionary saving motives. Finally, A_3^u is the temperance ratio[4].

It is easy to verify that if utility function u exhibits "mixed risk aversion", then the marginal utility function u' is completely monotone[5]. Thus, by Hausdorff-Bernstein-Widder's Theorem (see Bernstein [1928], Hausdorff [1921] and Widder [1931]), u' can be represented by

$$u'(x) = \int_0^\infty e^{-xs} dF_u(s),$$

[3]A function $f \in C(\mathbb{R}_+) \cap C^{k-2}(\mathbb{R}_{++})$, $k \geq 2$, for which $(-1)^l f^{(l)} \geq 0$, non-increasing and, convex for $l = 0, 1, 2, ..., k-2$ is called k-times monotone on $\mathbb{R}_{++}$. In the case $k = 1$, we only require f to be non-negative and non-increasing.

[4]Following Eeckhoudt, Gollier and Schneider (1995), an individual is said to be temperate, i.e. marginal gains in expected utility for successive upwards shifts of any increase in risk are decreasing. For temperate individuals, we have $A_3 > 0$.

[5]A $C(\mathbb{R}_+) \cap C^\infty(\mathbb{R}_{++})$ function f is said to be completely monotone if for all $k \in \mathbb{N}$ and $z \in \mathbb{R}_{++}$, we have that $(-1)^k f^{(k)}(z) \geq 0$.

where F_u is a distribution function on $\mathbb{R}_+$. This means that u' is the Laplace transform of F. It is well known that for all $n \in \mathbb{N}$ the limit, $\lim_{x \to 0^+} u^{(n)}(x)$ exists but is not necessary finite; more specifically, $\lim_{x \to 0^+} u^{(n)}(x)$ is finite if and only if $\int_0^\infty s^{n-1} dF_u(s) < \infty$. In addition, as shown in Caballé and Pomansky (1996), a representation of the utility function u is given by

$$u(x) = u(0) + \int_0^\infty \frac{1 - e^{-xs}}{s} dF_u(s),$$

where distribution F_u is such that $\int_1^\infty \frac{dF_u(s)}{s} < \infty$. Next, the following proposition provides a characterization of "mixed risk aversion" in terms of coefficient of absolute risk aversion of order n.

Proposition 1. *Assume that utility function u is $C(\mathbb{R}_+) \cap C^\infty(\mathbb{R}_{++})$, with $u' > 0$ such that for all $x \in \mathbb{R}_+$, $u^{(n+1)}(x) \neq 0$. Then, u exhibits "mixed risk aversion" if and only if for all $n \in \mathbb{N}$, A_n^u is a positive non-decreasing sequence.*

Proof. See the Appendix. □

Proposition 1 implies that A_n^u admits a limit (not necessarily finite) as n goes to ∞. Furthermore, for negative exponential utility functions, coefficient A_n^u is constant for all $n \in \mathbb{N}$ so even if $(-1)^n u^{(n+1)} > 0$ for all $n \in \mathbb{N}$, A_n^u may not be an increasing sequence. Finally, observe that A_n^u is a non-increasing function[6] of x since

$$\begin{aligned}
(A_n^u)'(x) &= -\frac{u^{(n+2)}(x)}{u^{(n)}(x)} + \left(\frac{u^{(n+1)}(x)}{u^{(n)}(x)} \right)^2 \\
&= -A_n^u(x) \left[A_{n+1}^u(x) - A_n^u(x) \right] \leq 0.
\end{aligned}$$

3.2.1 Higher Order Risk Changes and the Agency Problem

Using these definitions, we observe that in our agency model, the distribution of results shifts when agent's effort level a changes, and given that a is a continuous variable, we can use an extended version of Ekern's (1980) definition of increase (decrease) in risk for higher order risk changes in order to accommodate our problem to a differentiable setting with distributions of results conditional to agent's effort. For our particular purposes, the key element to have in mind is the partial derivative of the iterated integral $F^k(x \mid a)$, which is denoted by $F_a^k(x \mid a)$. As a result, FOSD will be associated with $F(x \mid a_2) - F(x \mid a_1) \leq 0$, with $a_1 \leq a_2$, which implies that in our continuous-effort moral hazard model, FOSD is equivalent to $F_a(x \mid a) \leq 0$, second order stochastic dominance (SOSD) will be associated

[6] The property $(A_n^u)' \leq 0$ was first derived by Caballé and Pomansky (1996).

with $F_a^2(x \mid a) \leq 0$, third order stochastic dominance (TOSD) will be associated with $F_a^3(x \mid a) \leq 0$ and so on. The following two definitions formalize these concepts.

Definition 5. *An increase in agent's effort represents a Nth degree stochastic (NSD) dominant shift if $F_a^N(x \mid a) \leq 0$ for all $x \in X$, where inequality is strict for some x, and $F_a^k(\overline{x} \mid a) \leq 0$ for $k = 1, 2, ..., N-1$.*

Definition 6. *An increase in agent's effort represents a Nth degree decrease in risk if $F_a^N(x \mid a) \leq 0$ for all $x \in X$ where inequality is strict for some x, and $F_a^k(\overline{x} \mid a) = 0$ for $k = 1, 2, ..., N-1$.*

The intuition behind the two definitions is straightforward when we use a discrete version of the effort model. Following Ekern (1980), let $\widetilde{x}_1$ and $\widetilde{x}_2$ be two random variables with a probability distribution given by $F(\cdot \mid a_1)$ and $F(\cdot \mid a_2)$ respectively, with $a_1 < a_2$. The condition $F_a^k(\overline{x} \mid a) = 0$ is equivalent to $F^k(\overline{x} \mid a_1) = F^k(\overline{x} \mid a_2)$ for all $k = 1, 2, ..., N-1$ in the discrete model. This means that the $N-1$ moments of both distributions are the same. For example, we can say that $\widetilde{x}_2$ is a second-degree decrease in risk of $\widetilde{x}_1$ if $\widetilde{x}_2$ dominates $\widetilde{x}_1$ via second order stochastic dominance and both distributions have equal mean, which is the case of the Rothschild and Stiglitz (1970) mean-preserving decrease in risk. A third-degree decrease in risk on the other hand, is equivalent to the concept of a decrease in downside risk developed by Menezes et al. (1980). Similarly, a fourth-degree decrease in risk is equivalent to the concept of a decrease in outer risk developed by Menezes and Wang (2005), and so on.

The last concept from the economics of risks literature that we will use in our paper is the concept of N^{th} degree stochastic change equivalence. Let V be a $C^N(X)$ utility function utility function. The expectation of V is defined as $E\left(V(\widetilde{x})\right) = \int_{\underline{x}}^{\overline{x}} V(x) dF(x \mid a)$ and after N iterations of integrating by parts, we find that

$$E\left(V(\widetilde{x})\right) = \sum_{k=1}^{N} (-1)^{k-1} V^{(k-1)}(\overline{x}) F^k(\overline{x} \mid a) + \int_{\underline{x}}^{\overline{x}} (-1)^N V^{(N)}(x) F^N(x \mid a) dx. \quad (7)$$

To evaluate the effect of a change in the agent's effort a on the expected utility, we differentiate (7) with respect to a to obtain

$$\int_{\underline{x}}^{\overline{x}} V(x) dF_a(x \mid a) = \sum_{k=1}^{N} (-1)^{k-1} V^{(k-1)}(\overline{x}) F_a^k(\overline{x} \mid a) + \int_{\underline{x}}^{\overline{x}} (-1)^N V^{(N)}(x) F_a^N(x \mid a) dx.$$

An increase in the agent's effort a that represents a NSD shift of the distribution F is equivalent to $\int_{\underline{x}}^{\overline{x}} V(x) dF_a(x \mid a) \geq 0$. Given what precedes, a sufficient condition for $\int_{\underline{x}}^{\overline{x}} V(x) dF_a(x \mid a) \geq 0$ is $(-1)^k V^{(k)}(x) \leq 0$ for all $x \in X$ and $k = 1, 2, ..., N$.

We conclude that if for all $n \leq N$, $(-1)^n V^{(n)}(x) \leq 0$ for all $x \in X$, an increase in agent's effort that causes N^{th}OSD shifts on $F(\cdot \mid a)$, the principal will always want the agent to exert a greater effort than what the latter is willing to provide under second-best sharing rule, i.e. $\mu > 0$.

4 Agency Theory and Stochastic Dominance

4.1 FOSD and Agency

The usual way that agency problems with continuous effort are solved is by using the first order condition given by Hölmstrom (1979). However, to be able to use Hölmstrom's conditions, we know, from the literature review section, that some mathematical properties of the functions involved in the problem must be satisfied, such the convexity of the distribution function and the monotone likelihood ratio condition (Mirrless (1977) and Rogerson (1985)), and given that we are using only general functional forms in this part of the paper, we will assume that all these technicalities are met and therefore, we can focus on the stochastic shifts of different order of the distribution of results and their effects on the optimal contract. For the FOSD case we could have referred the reader to the original source without any exposition here, but instead, we believe that it is important to provide an short version of Hölmstrom's proof in order to highlight the importance of the assumption of FOSD in the agency literature. Then, we will build on this result to analyze higher order stochastic dominant shifts.

Proposition 2. *Hölmstrom (1979)*
If an increase in agent's effort induces a FOSD shift of the distribution F, i.e., a shift that satisfies the property of $F_a^1(\cdot \mid a) \leq 0$, the principal would always desire a greater effort from the agent than the latter would be willing to provide under the second best sharing rule, i.e. $\mu > 0$.

Proof. See the Appendix. □

The previous proof relies on the concept of FOSD. But, what if the agent's effort affects the distribution of results through higher order stochastic dominance changes? FOSD is only one alternative of a stochastic change; it may even be the simplest way to think in the effect of agent's effort on the distribution of results. However, higher order stochastic changes can also make economic sense in other contexts as we will see below.

4.2 SOSD and Agency

As we discussed above, FOSD is a common and very reasonable assumption to represent the effects of agent's increasing effort on the distribution of results. However, in some cases, SOSD could be a better way to represent the effect of an agent's effort on a delegation set up under an agency problem. For instance, when a risk averse principal needs to delegate the management of a portfolio of new venture projects to an agent, it may be the case that the principal is very sensitive to the dispersion of a distribution and in consequence, SOSD shifts may provide the right preference ordering of distributions. Let us think in the case when a principal is about to retire (short-run horizon of investment) and he is not willing to take risky bets in the entrepreneurial world, the opposing interests between the principal and the agent come not only from the results (effort) - payment combination they have to agree upon, but also from the potentially different preferences of the risk-return trade off they may have. In this case, a very risk-averse principal may want to minimize risky decisions within the company portfolio of ventures, and the optimal contract has to be designed in a way that consider agent's effort cost and also his valuation of the risk-return trade off, which in this case would be decreasing dispersion. Alternatively, we can think in a classical agency problem within a firm where there exist a given hierarchy like in classical paper of Garicano (2000). In this context, we can find multiple reasons why a risk averse "boss" may want the agent to choose low risk tasks in order to minimize aggregate firm risk. Examples of these reasons are compensation schedules negotiated between the boss and the board that the boss do not want to put at risk, or avoiding the problem of reverse-delegation of risky and complex tasks within the organization when the boss can not observe directly the skill level of the agent and the latter also want to avoid being exposed as underperformed.

To the best of our knowledge, Hughes (1982) is the only application of agency theory under SOSD shifts induced by increasing agent's efforts. In this subsection, we generalize the results of Hughes (1982) that are based on Harmonic Absolute Risk Aversion (HARA) preferences, which are preferences that exhibit linear absolute risk tolerance. We use instead a general formulation of absolute risk tolerance to show that SOSD is the key property that defines the optimal contract when the principal is mainly interested in avoiding disperse distributions of outcomes. This analysis will be a way to introduce our general result for the case of Nth order stochastic dominance shifts induced from exerting more agent's effort.

Proposition 3. ***Generalizing Hughes (1982)***
If an increase in the agent's effort induces a SOSD shift of the distribution F, i.e. a shift that satisfies the property of $F_a^2(\cdot \mid a) \leq 0$, the principal would always want a greater effort

from the agent than the latter would be willing to provide under the second-best sharing rule, i.e. $\mu > 0$. *A sufficient condition is that for all* $x \in X$, $(V^*)''(x) \leq 0$, *which is satisfied whenever for all* $x \in X$, $(r^*)''(x) \leq 0$. *This last condition holds in the special case when the marginal risk tolerance of the agent is (always) higher than that of the principal.*

Proof. See the Appendix. □

Before examining in details necessary and sufficient conditions on the preferences of the principal and the agent for (i) $(V^*)''$ to be negative and (ii) $(r^*)''(x)$ to be negative, we present a sufficient condition the preferences of the principal and the agent that is easy to interpret.

Recall that the absolute risk tolerance T_1 is defined as the inverse of the absolute risk aversion coefficient that is $T_1 = \dfrac{1}{A_1}$. When the preferences are DARA, which is the case if the utility function exhibits the mixed risk aversion property, an increase in wealth implies a decrease in the absolute risk aversion coefficient, which in turns implies that $(T_1^u)' > 0$ and $(T_1^G)' > 0$. Differentiating with respect to x relationship (2) and rearranging terms, we have

$$(r^*)'(x) = \frac{T_1^G(r^*(x))}{T_1^u(s^*(x)) + T_1^G(r^*(x))} \in [0,1]. \tag{8}$$

Differentiating with respect to x relationship (8) and rearranging terms leads to

$$(r^*)''(x) = \frac{\left[(T_1^G)'(r^*(x)) - (T_1^u)'(s^*(x))\right](r^*)'(x)\left(1 - (r^*)'(x)\right)}{T_1^u(s^*(x)) + T_1^G(r^*(x))}. \tag{9}$$

Relationship (9) shows that, as long as $(T_1^G)'(r^*(x)) \leq (T_1^u)'(s^*(x))$, then $(r^*)''(x) \leq 0$. Evaluated at the optimal sharing rules, the marginal absolute tolerance of the agent must be greater than the marginal absolute risk tolerance of the principal. The intuition behind this result is clear. We would expect that the more marginally risk tolerant the agent is, the better he can cope with the risk of not having the best distribution of results (better than the principal), especially if the agent has to assume all the effort's cost to obtain second-order stochastic dominant distributions. In other words, the agent's second-best optimal effort is less than what the principal would like him to exert when the agent's marginal absolute risk tolerance is higher than the principal's.

Clearly, if there exists some $\alpha \in \mathbb{R}$ such for all $(y,z) \in \Delta_G \times \Delta_u$, $(T_1^G)'(y) \leq \alpha \leq (T_1^u)'(z)$, which is equivalent to

$$\frac{A_2^G(y)}{A_1^G(y)} \leq \alpha \leq \frac{A_2^u(z)}{A_1^u(z)}, \text{ for all } (y,z) \in \Delta_G \times \Delta_u,$$

then we have $(r^*)''(x) \leq 0$. This condition is quite strong as it imposes a uniform property on coefficients A_k^G and A_k^u for $k = 1, 2$. In section 6 of the paper, we shall provide a weaker (necessary and sufficient) condition for the class of HARA utility functions thereby generalizing the results of Hughes (1982).

In the next section, we investigate the concavity of functions V^* and r^* and look for weaker conditions.

4.2.1 Study of $(V^*)''$

Recall that we have established that $r^* = I_{\Gamma_u}$ where function I_{Γ_u} is given by relationship (3) so that for $x \in X$, $V^*(x) = G(I_{\Gamma_u}(x))$. Differentiating with respect to x leads to

$$\begin{aligned} (V^*)'(x) &= I'_{\Gamma_u}(x) G'(I_{\Gamma_u}(x)) > 0 \\ (V^*)''(x) &= I''_{\Gamma_u}(x) G'(I_{\Gamma_u}(x)) + (I'_{\Gamma_u}(x))^2 G''(I_{\Gamma_u}(x)). \end{aligned}$$

We start by providing a condition so that $I''_{\Gamma_u}(x) = (r^*)''(x) \leq 0$ for all $x \in X$.

Necessary and Sufficient Condition for $(r^*)'' \leq 0$. Let I_G and I_u denote the inverse functions of the G' and u' respectively. Note that $u'(I_u(z)) = z$, so we have

$$I'_u(z) = \frac{1}{u''(I_u(z))} < 0 \text{ and } I''_u(z) = -\frac{u'''(I_u(z))}{[u''(I_u(z))]^3} > 0 \text{ (assuming } u''' > 0).$$

As $r^*(x) = I_{\Gamma_u}(x)$, differentiating with respect to x yields

$$\begin{aligned} (r^*)'(x) &= I'_{\Gamma_u}(x) = \frac{1}{\Gamma'_u(I_{\Gamma_u}(x))} > 0. \\ (r^*)''(x) &= I''_{\Gamma_u}(x) = -\frac{\Gamma''_u(I_{\Gamma_u}(x))}{[\Gamma'_u(I_{\Gamma_u}(x))]^3}. \end{aligned}$$

Hence $(r^*)'' \leq 0$ if and only if $\Gamma''_u \geq 0$. Note that

$$\Gamma'_u(y) = \frac{I'_G(G'(y)) + I'_u(G'(y))}{I'_G(G'(y))},$$

and

$$\Gamma''_u(y) = \frac{I'_u(G'(y))}{[I'_G(G'(y))]^2} \left[A_1^{I_G}(G'(y)) - A_1^{I_u}(G'(y)) \right],$$

where, $A_1^{I_u} = -\frac{I''_u}{I'_u}$, $A_1^{I_G} = -\frac{I''_G}{I'_G}$. Since $I'_u < 0$ and also $I'_G < 0$, it follows that $(r^*)'' \leq 0$ if and only if

$$A_1^{I_G}(z) \leq A_1^{I_u}(z) \text{ for all } z \text{ in } ImG'. \tag{10}$$

Observe that the relationship has to hold only on the set ImG', i.e., for z such that there exists y in X such that $z = G'(y)$.

We now provide a general necessary and sufficient condition on utility functions u and G so that V^* is concave.

Necessary and Sufficient Condition for $(V^*)'' \leq 0$. From the expressions derived for $(V^*)''$, $I'_{\Gamma_u}(x)$ and $I''_{\Gamma_u}(x)$ we have

$$(V^*)''(x) = \frac{G'(I_{\Gamma_u}(x))}{[\Gamma'_u(I_{\Gamma_u}(x))]^2}\left[\frac{G''(I_{\Gamma_u}(x))}{G'(I_{\Gamma_u}(x))} - \frac{\Gamma''_u(I_{\Gamma_u}(x))}{\Gamma'_u(I_{\Gamma_u}(x))}\right].$$

Set $z = G'(I_{\Gamma_u}(x)) = G'(r^*(x))$ so that $I_G(z) = I_{\Gamma_u}(x)$ and observe that

$$\begin{aligned} G''(I_{\Gamma_u}(x)) &= \frac{1}{I'_G(z)} \\ \Gamma'_u(I_{\Gamma_u}(x)) &= \frac{I'_u(z) + I'_G(z)}{I'_G(z)}. \end{aligned}$$

It follows that

$$(V^*)''(x) = \frac{I'_G(z)}{[I'_u(z) + I'_G(z)]^2}\left[1 + z\frac{I'_u(z)}{I'_G(z) + I'_u(z)}\left[A_1^{I_u}(z) - A_1^{I_G}(z)\right]\right].$$

As $I'_G < 0$ we have $(V^*)''(x) \leq 0$ if and only if

$$1 + z\frac{I'_u(z)}{I'_G(z) + I'_u(z)}\left[A_1^{I_u}(z) - A_1^{I_G}(z)\right] \geq 0 \text{ for all } z \text{ in } ImG'. \tag{11}$$

Note that condition (11) involves the inverse of the marginal utility functions and their coefficients of risk aversion instead of the marginal utilities themselves.

In the next section, we generalize our analysis to the case of Nth order stochastic dominance shifts.

4.3 $N^{\mathbf{th}}$OSD and Agency

In the previous sections we have discussed specific conditions under which different possible types of stochastic dominant shifts in the distribution of results occur when the agent exerts more effort. We have provided some economic intuition for the cases of FOSD, SOSD and TOSD. Analyzing higher order stochastic dominant shifts turns out to be a much more challenging task. For example, economic intuition for FOOSD is likely to be connected with aversion to outer risk (Menezes and Wang, 2005), but the algebraic tractability of the proof becomes cumbersome.

For increases in risk, Diamond and Stiglitz (1974) characterize the role of greater risk aversion by considering the class of distributional changes in risk that induce mean preserving spreads (or alternatively contractions) in the probability distribution of results. The condition for the utility function is straightforward, v is more risk averse than u if and only if v is a concave transformation of u, i.e., $v = \varphi(u)$ where φ is an increasing concave function (Pratt [1964]). A TOSD shift can also be characterized by altering the notion of a mean-preserving spreads (Rothschild and Stiglitz (1970)) into a variance-preserving third-order risk spread, and then showing that a downside risk-averse person is indeed characterized by a dislike of all such spreads. Menezes, Geiss and Tressler (1980) show that downside risk aversion is characterized by $u''' \geq 0$. Keenan and Snow (2002) show that a measure of downside risk aversion is $s_u = \frac{u'''}{u'} - \frac{3}{2}\left(\frac{u''}{u'}\right)^2 = \frac{A_2^u}{A_1^u} - \frac{3}{2}(A_1^u)^2$ in the case of small increases in downside risk aversion.

Recall that $V(x) \equiv G\left(r^*\left(x\right)\right)$ and observe that when the principal is N^{th} order risk averse, we have $(-1)^k G^{(k)} \leq 0$ for $k = 1, ..., N$. Aforementioned, an increase in agent's effort a that represents a NOSD shift on $F\left(\cdot \mid a\right)$ is equivalent to $\int_{\underline{x}}^{\overline{x}} V(x) dF_a(x \mid a) \geq 0$, which is satisfied if $(-1)^k V^{(k)} \leq 0$ for $k = 1, ..., N$.

4.3.1 General Condition for $(-1)^n (V^*)^{(n)} \leq 0$

We wish to provide a sufficient and necessary condition on the coefficients $A_k^{I_u}$ and $A_k^{I_G}$ with $k = 1, ..., n-1$ so that V^* satisfies the mixed risk aversion property up to order n. Working with the inverses of the marginal utility functions instead of working directly with utility functions allows us to derive a condition on a functional that can be evaluated at some common argument[7], namely variable $z = G'(r^*(x))$. Given the assumptions on utility functions G and u, for $k = 1$, coefficients $A_1^{I_G}$ and $A_1^{I_u}$ are well defined and positive for all x in X. However, for k integer greater or equal to 2, we may have (for instance) $I_G^{(k+1)}(x) = 0$ for some x in X with $I_G^{(k)}(x) \neq 0$. In this case, we have $A_k^{I_G}(x) = 0$ but $A_{k+1}^{I_G}(x)$ is not well defined. Then, to move from order k to order $k+1$, we need to be able to compute quantity $(A_n^{I_G})'(x)$. Even if $A_{k+1}^{I_G}(x)$ is not well defined, the "symbolic" product $A_{k+1}^{I_G}(x) A_k^{I_G}(x)$ is equal to $-\frac{I_G^{(k+2)}(x)}{I_G^{(k)}(x)}$ so it is well defined. Therefore, we can still write - symbolically - $(A_n^{I_G})'(x) = -A_n^{I_G}(x)\left[A_{n+1}^{I_G}(x) - A_n^{I_G}(x)\right]$.

For $n = 1$, we have

$$\begin{aligned} (V^*)'(x) &= z \frac{I_G'(z)}{I_G'(z) + I_u'(z)}, \\ &= Q_1(z, I_G'(z), I_u'(z)), \end{aligned}$$

[7]Unfortunately, this is not the case if one intends to derive a condition that involves coefficients A_k^u and A_k^G.

where $Q_1(X_0, X_1, X_2) = X_0 \frac{X_1}{X_1+X_2}$. For $n = 2$, we have seen that for all $x \in X$, $(V^*)''$ can be expressed as

$$(V^*)''(x) = Q_2(z, I'_G(z), I'_u(z), A_1^{I_G}(z), A_1^{I_u}(z)), \tag{12}$$

where Q_2 is a rational function of $2 \times 2 + 1 = 5$ variables that is given by

$$Q_2(X_0, X_1, X_2, X_3, X_4) = \frac{X_1}{(X_1 + X_2)^2} \left[1 + X_0 \frac{X_2}{X_1 + X_2} [X_4 - X_3] \right].$$

Then, observe that $(G'(r^*(x))' = \frac{1}{I'_u(z)+I'_G(z)}$, $I''_u(z) = -A_1^{I_u}(z) I'_u(z)$, $I''_G(z) = -A_1^{I_G}(z) I'_G(z)$ and recall that

$$\begin{aligned} (A_n^{I_G})'(z) &= -A_n^{I_G}(z) \left[A_{n+1}^{I_G}(z) - A_n^{I_G}(z) \right] \\ (A_n^{I_u})'(z) &= -A_n^{I_u}(z) \left[A_{n+1}^{I_u}(z) - A_n^{I_u}(z) \right]. \end{aligned}$$

Therefore, given what precedes, differentiating relationship (12) with respect to x, it is possible to write

$$(V^*)'''(x) = Q_3(z, I'_G(z), I'_u(z), A_1^{I_G}(z), A_1^{I_u}(z), A_2^{I_G}(z), A_2^{I_u}(z)),$$

with $z = G'(r^*(x))$ and where Q_3 is a rational function of $2 \times 3 + 1 = 7$ variables. More generally, we have the following proposition.

Proposition 4. *The derivative of order n of function V^* can be expressed as a rational function of $2n + 1$ variables $(X_0, X_1, ..., X_{2n})$; variables X_0, X_1 and X_2 that are to be evaluated at z, $I'_G(z)$ and $I'_u(z)$ respectively whereas for $k \geq 1$, variables x_{2k+1} and x_{2k+2} are to be evaluated at $A_k^{I_G}(z)$ and $A_k^{I_u}(z)$ respectively, with $z = G'(r^*(x))$. Furthermore, functionals Q_n and Q_{n+1} are linked by the following relationship*

$$\begin{aligned} Q_{n+1}(X_0, X_1, ..., X_{2n+2}) \;=\; & \frac{1}{X_1 + X_2} \left[\frac{\partial Q_n}{\partial X_0} - X_1 X_3 \frac{\partial Q_n}{\partial X_1} - X_2 X_4 \frac{\partial Q_n}{\partial X_2} \right. \\ & \left. - \sum_{k=1}^{n-1} X_{2k+1}(X_{2k+3} - X_{2k+1}) \frac{\partial Q_n}{\partial X_{2k+1}} - \sum_{k=1}^{n-1} X_{2k+2}(X_{2k+4} - X_{2k+2}) \frac{\partial Q_n}{\partial X_{2k+2}} \right] \end{aligned}$$

Proof. The property is true for $n = 1$. Then, for $n \geq 1$, one can check that the induction relationship of the proposition holds. □

Function $(V^*)^{(n)}$ depends of all the coefficients of risk aversion of functions I_G and I_u up to order $n - 1$, then for large n, it may be cumbersome to infer its sign. In the remainder of the paper, we look for sufficient conditions so that V^* satisfies the mixed risk aversion property.

4.3.2 Sufficient Condition for $(-1)^n (V^*)^{(n)} \leq 0$

Proposition 5. *Assume that functions G and r^* are C^N differentiable functions that satisfy*

$$\begin{aligned}(-1)^n G^{(n)}(x) &\leq 0 \\ (-1)^n (r^*)^{(n)}(x) &\leq 0, \text{ for all } x \in X \text{ and all integers } n \leq N.\end{aligned}$$

Then function V^* satisfies the mixed risk aversion property, i.e., for all $x \in X$ and all integers $n \leq N$, we have

$$(-1)^n (V^*)^{(n)}(x) \leq 0.$$

Proof. See the Appendix. □

4.3.3 A Sufficient Condition for $(-1)^n (r^*)^{(n)} \leq 0$

Recall that we have established that

$$\Gamma_u(r^*(x)) = x,$$

with $\Gamma_u(y) = y + I_u(G'(y))$. We now present a sufficient condition on function Γ_u such that its inverse satisfies the mixed risk aversion property. We start with the following lemma.

Lemma 1. *Let f be a function in $C(\mathbb{R}_+) \cap C^\infty(\mathbb{R}_{++})$ such that $f' > 0$ and for all integers $k \geq 2$ and $z > 0$, $(-1)^k f^{(k)}(z) \geq 0$. Then, the inverse function of f satisfies the mixed risk aversion property.*

Proof. See the Appendix. □

We apply the result of lemma 1 to function Γ_u whose derivative is strictly positive. We conclude that a sufficient condition for r^* to satisfy the mixed risk aversion property is function Γ_u' to be completely monotone. Clearly, this last condition is easier to verify.

We now illustrate our results for standard utility functions used in the financial economics literature.

5 Examples

In order to lighten the notation, we use s and r to refer to the optimal Pareto rules instead of s^* and r^* and V instead of V^*. Finally, recall that we normalize λ^* to be equal to 1.

5.1 CARA Utility Functions

Let $G(x) = \frac{1-e^{-Bx}}{B}$ and $u(x) = \frac{1-e^{-Ax}}{A}$, with A and B positive constants. For this class of utility functions, for the principal (resp. agent), we have $A_n(x) = B$ (resp. A); in particular, the absolute tolerance is constant, so the marginal absolute tolerance is equal to zero. Furthermore, it is easy to verify that G' and u' are the Laplace transforms of the distributions $F_G(s) = H(s-B)$ and $F_u(s) = H(s-A)$ respectively, where H denotes the Heaviside function. In this example, we can choose $X = \mathbb{R}$.

Relationship (2) can be written

$$\frac{G'(x-s(x))}{u'(s(x))} = e^{-Bx+(A+B)s(x)} = 1 > 0,$$

which leads to a linear sharing rule[8]

$$s(x) = \frac{B}{A+B}x.$$

We find that for all $x \in \mathbb{R}$, $V(x) = \dfrac{1-e^{-\frac{AB}{A+B}x}}{B}$ and therefore the property $(-1)^k V^{(k)}(x) \leq 0$ is satisfied for all $k \in \mathbb{N}$.

This simple example illustrates that no restrictions are needed for function V to satisfy the mixed risk aversion property.

5.2 CRRA Utility Functions

Let $G(x) = \frac{x^{1-B}-1}{1-B}$ (resp. $\ln x$) if $B \neq 1$ (resp. $B = 1$) and $u(x) = \frac{x^{1-A}-1}{1-A}$ (resp. $\ln x$) if $A \neq 1$ (resp. $A = 1$) , with A and B positive constants. In this case, we choose $X = \mathbb{R}_+$. For this class of utility functions, we have $A_n^G(x) = \frac{B+n-1}{x}$ and $A_n^u(x) = \frac{A+n-1}{x}$; in particular, the absolute tolerance is linear, so the marginal absolute tolerance is constant. Furthermore, it is easy to verify that G' and u' are the Laplace transforms of the distributions $F_G(s) = \frac{s^B}{\Gamma(B+1)}$ and $F_u(s) = \frac{s^A}{\Gamma(A+1)}$ respectively, where Γ denotes Euler's Gamma function. Relationship (2) can be written

$$(x - s(x))^B = (s(x))^A, \tag{13}$$

for all $x \in \mathbb{R}_+$ and $s(x) \in [0, x]$. Then, we have $I_G(z) = z^{-\frac{1}{B}}$, $I_u(z) = z^{-\frac{1}{A}}$ so that $A_1^{I_G}(z) = \frac{1+B}{B}\frac{1}{z}$ and $A_1^{I_u}(z) = \frac{1+A}{A}\frac{1}{z}$. Given conditions (10) and (11) we have $r'' \leq 0$ if and only if $A \leq B$ and $V'' \leq 0$ if and only if $\frac{A}{1+A} \leq B$.

[8] More generally, it is easy to verify that the optimal payoff sharing rule is linear if utility functions G and u are such that $G(x) = \frac{1}{a}u(ax) + b$ for all $x \in X$ with constants $(a, b) \in \mathbb{R}_{++} \times \mathbb{R}$.

5.2.1 Case $A = 1$ and $B = 2$

A closed form solution for s and r is available and it is easy to verify that

$$\begin{aligned} s(x) &= \frac{1 + 2x - \sqrt{1 + 4x}}{2} \\ r(x) &= \frac{\sqrt{1 + 4x} - 1}{2}. \end{aligned}$$

Thus, for all integers $n \geq 2$

$$s^{(n)}(x) = (-1)^n \times 1 \times 3 \times ... \times (2n - 3)2^{n-1}(1 + 4x)^{-\frac{(2n-1)}{2}}.$$

Hence we have

$$\begin{aligned} (-1)^{n+1}s^{(n)}(x) &< 0 \\ (-1)^n r^{(n)}(x) &< 0. \end{aligned}$$

We can claim that for all $x \in \mathbb{R}_{++}$ and $n \in \mathbb{N}$, $(-1)^n V^{(n)}(x) \leq 0$.

5.2.2 Case $A = 2$ and $B = 1$

Again, a closed form solution is available and it is easy to verify that

$$\begin{aligned} s(x) &= \frac{-1 + \sqrt{1 + 4x}}{2} \\ r(x) &= \frac{2x + 1 - \sqrt{1 + 4x}}{2}. \end{aligned}$$

It follows that for all integers $n \geq 2$

$$s^{(n)}(x) = (-1)^{n+1} \times 1 \times 3 \times ... \times (2n - 3)2^{n-1}(1 + 4x)^{-\frac{(2n-1)}{2}}.$$

Hence we have

$$\begin{aligned} (-1)^{n+1}s^{(n)}(x) &> 0 \\ (-1)^n r^{(n)}(x) &> 0. \end{aligned}$$

Even though we have $r'' > 0$, observe that as $\frac{A}{1+A} = \frac{2}{3} < B = 1$, we still have $V'' < 0$.

5.2.3 General Case

We have already seen that $r'' < 0$ if and only if $B > A$. Thus, we can restrict our attention to the case $B > A$. Totally differentiating relationship (13) with respect to x and rearranging terms leads to

$$s'(x) = \frac{Bs(x)}{A(x - s(x)) + Bs(x)} \in (0, 1).$$

Differentiating one more time yields

$$s''(x) = (B - A)AB \frac{s(x)(x - s(x))}{[A(x - s(x)) + Bs(x)]^3}.$$

Differentiating one more time with respect to x after rearranging terms we find that

$$s'''(x) = \frac{(s''(x))^2 x}{(B - A)s'(x)(1 - s'(x))} \varphi(x),$$

with $\varphi(x) = -2A^2(1 - \varpi(x)) - 2B^2\varpi(x) + AB$ and $\varpi(x) = \frac{s(x)}{x} \in (0, 1)$. Next, note that for all $x > 0$, $\varpi(x)$ is implicitly defined by

$$x^{B-A}(1 - \varpi(x))^B = (\varpi(x))^A.$$

It is easy to verify that when $B > A$, ϖ is a strictly increasing function with $\lim_{0^+} \varpi = 0$ and $\lim_{\infty} \varpi = 1$. Clearly, when $B > A$, function φ is decreasing in x with $\varphi(0) = A(B-2A)$ and $\lim_{\infty} \varphi = B(A - 2B) < 0$. We conclude that φ has a constant sign, namely negative, if and only if $A < B \leq 2A$. To sum up, $(-1)^3 r'''(x) < 0$ for all $x \geq 0$ iff $B \leq 2A$ and in general, s''' does not have a constant sign. We now provide a full characterization for all order n.

Proposition 6. *For all $x \in \mathbb{R}_{++}$ and $n \in \mathbb{N}$, we have*

$$(-1)^n r^{(n)}(x) \leq 0 \text{ if and only if } A \leq B \leq 2A.$$

Proof. See the Appendix. □

5.3 HARA Utility Functions

For $a \neq 0$, $b \neq 0$, $A > 0$ and $B > 0$, let $G(x) = \frac{B}{1-B}(b + \frac{x}{B})^{1-B}$ and $u(x) = \frac{A}{1-A}(a + \frac{x}{A})^{1-A}$, so that $T_1^G(x) = b + \frac{x}{B}$ and $T_1^u(x) = a + \frac{x}{A}$ and we assume that $x > -\max\{-aA, -bB\}$. It follows that $I_G(z) = Bz^{-1/B} - bB$ and $I_u(z) = Az^{-1/A} - aA$. Thus, we have $I'_G(z) = -z^{-1/B-1}$, $I'_u(z) = -z^{-1/A-1}$, $A_n^{I_G}(z) = \frac{B+n}{Bz}$ and $A_n^{I_u}(z) = \frac{A+n}{Az}$. These expressions are independent of parameters a and b and therefore the results derived for the class of CRRA utility functions still apply.

We note that we have $V'' < 0$ if and only if $\frac{A}{1+A} \leq B$, which generalizes the condition derived by Hughes (1982), namely $(T_1^G)' < (T_1^u)'$, i.e., $A \leq B$.

5.4 $V'' = 0$

In this section, we look for a sufficient and necessary condition for $V'' = 0$ and provide an example of preferences that satisfy this condition. Obviously, the property $(-1)^n V^{(n)} \leq 0$ will be satisfied.

If for all $x \in X$, $V''(x) = 0$ then we have $V'(x) = K > 0$ for all $x \in X$. It follows that

$$G'(I_{\Gamma_u}(x)) = \frac{K}{I'_{\Gamma_u}(x)} = K\Gamma'(I_{\Gamma_u}(x)),$$

or equivalently

$$\frac{z-K}{K} I'_G(z) = I'_u(z). \tag{14}$$

where $z = I_{\Gamma_u}(x)$. Since $I'_u < 0$ and $I'_G < 0$, we must impose $z > K > 0$. Note that condition (14) implies that

$$A_1^{I_G}(z) = A_1^{I_u}(z) + \frac{1}{z-K}.$$

Assume that $u(x) = -\frac{e^{-ax}}{a}$ so that $T_1^u(x) = \frac{1}{a}$. It follows that $I'_u(z) = -\frac{1}{az}$ and thus

$$I'_G(z) = \frac{1}{a}\left[\frac{1}{z} - \frac{1}{z-K}\right].$$

Integrating this equation yields

$$I_G(z) = \frac{1}{a}\ln\frac{z}{z-K} + c,$$

for some constant c. We may choose $c \geq 0$, so that I_G is well defined on (K, ∞) and define set $X = (c, \infty)$. Inverting the relationship leads to

$$G'(x) = \frac{K}{1 - e^{-a(x-c)}},$$

for all $x \in X$ so that

$$G(x) = Kx + \frac{K}{a}\ln\left[1 - e^{-a(x-c)}\right] + C,$$

where C is a constant. Utility function G is of the DARA type with $T_1^G(x) = \frac{1}{a}\left[e^{a(x-c)} - 1\right]$; note that $(T_1^G)'(x) > (T_1^u)'(x) = 0$, so the condition $(T_1^G)' < (T_1^u)'$ does not hold. Furthermore, one can check that the optimal payoff sharing rule for the manager is given by

$$r(x) = c + \frac{1}{a}\ln\left[1 + Ke^{a(x-c)}\right],$$

where $x \in X$. Observe that

$$\begin{aligned} r'(x) &= \frac{Ke^{a(x-c)}}{1+Ke^{a(x-c)}} > 0 \\ r''(x) &= \frac{Kae^{a(x-c)}}{\left[1+Ke^{a(x-c)}\right]^2} > 0, \end{aligned}$$

and therefore (not surprisingly) we do not have $r'' \leq 0$.

6 Conclusion

In this paper we build a bridge between recent developments in the economic theory of risk and uncertainty and the classic agency problem. We derive conditions on the agent and the principal preferences under which if the effects of the agent's efforts are represented by a Nth-order stochastic dominance the principal would want a greater effort level than the one provided under the second best sharing rule. This allows us to extend the agency analysis to a variety of economic settings beyond the usual first-order stochastic dominance paradigm. We argue that second-order or third-order stochastic dominance may provide a suitable ordering of prospects. For instance, in the case of delegated entrepreneurial investment decisions, a principal may want the agent to minimize risk (second-order risk decision) or alternatively, the principal may want the agent to avoid losses of any kind (downside risk aversion, prudent behavior or third-order risk decision). In these cases, first-order stochastic dominance is not the relevant criteria to rank the distributions of outcomes and consequently, it is important to consider higher-order risk attitudes to analyze more general moral hazard issues and implications.

Our characterization involves the inverses of the marginal utilities of the agent and the principal as well as their derivatives up to order N. One challenge is that the properties on the utilities functions and their derivatives only transmit in an increasing cumbersome way to the derivatives of the inverse marginal utilities as the order of the derivatives rises. When the principal preferences exhibit mixed risk aversion, a simplified approach examines sufficient (but easier to verify) conditions and highlights the mixed risk aversion property that must satisfy the principal optimal Pareto share rule. Finally, we illustrate our findings for standard preferences used in financial economics.

This paper is among the first attempts to incorporate more complex risk structures into traditional economic models. Many challenges remain but we believe this is an important step towards a more acute understanding of seminal or new economic frameworks.

7 Appendix

Appendix A

Proof of Proposition 1. **Step 1** ($\Rightarrow$): Assume that $(-1)^n u^{(n+1)} > 0$ so that for all $x \in \mathbb{R}_{++}$ we can write

$$u'(x) = \int_0^\infty e^{-xs} dF_u(s),$$

for some distribution function F_u. It follows that for all $n \in \mathbb{N}$, $u^{(n)}(x) = (-1)^{n-1} \int_0^\infty s^{n-1} e^{-xs} dF_u(s)$, and therefore

$$A_n^u(x) = \frac{\int_0^\infty s^n e^{-xs} dF_u(s)}{\int_0^\infty s^{n-1} e^{-xs} dF_u(s)}.$$

To show that $A_{n+1}^u \geq A_n^u$ it is equivalent to prove that

$$\int_0^\infty s^{n+1} e^{-xs} dF_u(s) \times \int_0^\infty s^{n-1} e^{-xs} dF_u(s) \geq \left(\int_0^\infty s^n e^{-xs} dF_u(s) \right)^2.$$

Consider the quadratic Q in ξ

$$Q(\xi) = \int_0^\infty (\xi s + 1)^2 s^{n-1} e^{-xs} dF_u(s).$$

Clearly, Q is a non-positive function so at most it has a double real root; therefore, its discriminant Δ must be non-positive, i.e.,

$$4 \left(\int_0^\infty s^n e^{-xs} dF_u(s) \right)^2 - 4 \int_0^\infty s^{n+1} e^{-xs} dF_u(s) \times \int_0^\infty s^{n-1} e^{-xs} dF_u(s) \leq 0.$$

This concludes step 1.

Step 2 ($\Leftarrow$): Conversely, assume that A_n is strictly increasing and positive. For $n = 1$, we have

$$A_1^u = -\frac{u''}{u'} > 0,$$

so $u'' < 0$ as $u' > 0$ (by assumption) and for $n = 2$, we have

$$-\frac{u'''}{u''} > -\frac{u''}{u'} > 0,$$

so we must have $u''' > 0$. We now show the result by induction on n. Assume that $(-1)^n u^{(n+1)} > 0$. We write

$$A_{n+1}^u = -\frac{u^{(n+2)}}{u^{(n+1)}} = \frac{(-1)^{n+1} u^{(n+2)}}{(-1)^n u^{(n+1)}} \geq A_n^u > 0.$$

As $(-1)^n u^{(n+1)} > 0$ we deduce that $(-1)^{n+1} u^{(n+2)} > 0$, which concludes step 2 and the proof is complete.

Appendix B

Proof of Proposition 2. By contradiction. Assume that $\mu < 0$. Let us denote s_λ^* (resp. r_λ^*) the Pareto sharing rule for the agent (resp. principal) corresponding to parameter λ and observe that by relationship (2), s_λ^* and r_λ^* are increasing in x. Then, when $\mu < 0$ relationship (4) guarantees that $s^{**}(x) \leq s_\lambda^*(x)$ on the set $\{x,\ f_a(x \mid a) \geq 0\}$, and $s^{**}(x) > s_\lambda^*(x)$ on the set $\{x,\ f_a(x \mid a) < 0\}$. Set $V^{**}(x) = G(r^{**}(x))$ and $V_\lambda^*(x) = G(r_\lambda^*(x))$. Consequently, we have

$$\int_X V^{**}(x) f_a(x \mid a)\, dx \geq \int_X V_\lambda^*(x) f_a(x \mid a)\, dx. \tag{15}$$

Integrating by parts the RHS of inequality (15) implies

$$\int_X V_\lambda^*(x) f_a(x \mid a)\, dx = -\int_X (V_\lambda^*)'(x) F_a(x \mid a)\, dx, \tag{16}$$

since $F_a(\underline{x} \mid a) = F_a(\overline{x} \mid a) = 0$ for all a. Then, for $\mu < 0$ by relationship (5), we have $\int_X V^{**}(x) f_a(x \mid a)\, dx < 0$, which implies that using relationship (15)

$$\int_X (V_\lambda^*)'(x) F_a(x \mid a)\, dx > 0, \text{ for } \mu < 0.$$

We know that $(V_\lambda^*)'(x) = G'(r_\lambda^*(x))(r_\lambda^*)'(x)$ and $G' > 0$. Then, differentiating with respect to x relationship (2) and after rearranging terms we obtain that

$$(r_\lambda^*)'(x) = \frac{A_1^u(r_\lambda^*(x))}{A_1^u(r_\lambda^*(x)) + A_1^G(s_\lambda^*(x))} > 0. \tag{17}$$

where A_1^u and A_1^G denote the Arrow-Pratt absolute risk aversion coefficients of the agent and principal respectively. Therefore, having established that $(r_\lambda^*)'(x) > 0$, since $G'(r_\lambda^*(x)) > 0$ and by FOSD we know that $F_a(x \mid a) < 0$, we conclude that $\int_X (V_\lambda^*)'(x) F_a(x \mid a)\, dx < 0$, which is a contradiction.

Appendix C

Proof of Proposition 3. The proof is similar the one provided for Proposition 2. We normalize λ to 1 and to lighten the notation we drop subscript λ so that $V_\lambda^* = V^*$ and $r_\lambda^* = r^*$. Integrating by parts the expression $-\int_X (V^*)'(x) F_a(x \mid a)\, dx$ we obtain

$$\begin{aligned}\textstyle\int_X V^*(x) f_a(x \mid a)\, dx &= -\textstyle\int_X (V^*)'(x) F_a(x \mid a)\, dx = \\ -(V^*)'(x) F_a^2(x \mid a) \Big|_{\underline{x}}^{\overline{x}} &+ \textstyle\int_X (V^*)''(x) F_a^2(x \mid a) dx.\end{aligned} \tag{18}$$

Then, relationship (18) becomes

$$\int_X V^*(x) f_a(x \mid a)\, dx = -(V^*)'(\overline{x}) F_a^2(\overline{x} \mid a) + \int_X (V^*)''(x) F_a^2(y \mid a)\, dx. \tag{19}$$

The definition of SOSD ensures that $F_a^2(\overline{x} \mid a) \leq 0$ and $F_a^2(\cdot \mid a) \leq 0$, and given that $(V^*)'(\overline{x})$ is positive, it is easy to see that a sufficient condition for $\int_X V^*(x) f_a(x \mid a)\, dx \geq 0$ is $(V^*)''(x) \leq 0$ for all $x \in X$. Recall that $V^*(x) = G(r^*(x))$, so differentiating twice with respect to x leads to

$$(V^*)''(x) = G''(r^*(x)) \cdot [(r^*)'(x)]^2 + G'(r^*(x)) \cdot (r^*)''(x).$$

Since $[(r^*)']^2 > 0$, $G' > 0$ and $G'' < 0$, $(r^*)''(x) < 0$ for all $x \in X$ ensures that $(V^*)'' < 0$ on X.

Appendix D

Proof of Proposition 6. The proof relies on the Faà di Bruno formula[9] (1857) that states that for function $V^* = G(r^*(x))$, for all $n \leq N$ we have:

$$(V^*)^{(n)}(x) = \sum \frac{n!}{b_1!...b_n!} G^{(j)}(r^*(x)) \prod_{i=1}^{n} \left(\frac{(r^*)^{(i)}(x)}{i!} \right)^{b_i},$$

with $b_i \in \mathbb{N}$ and

$$\begin{aligned} b_1 + ... + b_n &= j \\ b_1 + 2b_2 ... + nb_n &= n. \end{aligned}$$

Let assume for now that for all $x \in X$ and all $i \leq N$, $(-1)^i (r^*)^{(i)}(x) \leq 0$. It follows that

$$\begin{aligned} (V^*)^{(n)}(x) &= \sum \frac{n!}{b_1!...b_n!} G^{(j)}(r^*(x)) \prod_{i=1}^{n} \left(\frac{(-1)^{i+1}(r^*)^{(i)}(x)}{i!} (-1)^{i+1} \right)^{b_i} \\ &= \sum \frac{n!}{b_1!...b_n!} G^{(j)}(r^*(x)) \prod_{i=1}^{n} \left(\frac{(-1)^{i+1}(r^*)^{(i)}(x)}{i!} \right)^{b_i} \prod_{i=1}^{n} [(-1)]^{(i+1)b_i} \\ &= \sum \frac{n!}{b_1!...b_n!} G^{(j)}(r^*(x)) \prod_{i=1}^{n} \left(\frac{(-1)^{i+1}(r^*)^{(i)}(x)}{i!} \right)^{b_i} [(-1)]^{j+n}, \end{aligned}$$

using the conditions on integers b_i. Thus

$$(V^*)^{(n)}(x) = (-1)^n \sum \frac{n!}{b_1!...b_n!} (-1)^j G^{(j)}(r^*(x)) \prod_{i=1}^{n} \left(\frac{(-1)^{i+1}(r^*)^{(i)}(x)}{i!} \right)^{b_i}.$$

Since by assumption $(-1)^{i+1}(r^*)^{(i)}(x) \geq 0$ and $(-1)^j G^{(j)}(r^*(x)) \leq 0$ for all $j \leq N$, we deduce $(-1)^n (V^*)^{(n)}(x) \leq 0$ for all $x \in X$.

[9]The Faà di Bruno formula (1857) provides an explicit expression the derivative of order n of a composite function and generalizes the chain rule.

Appendix E

Proof of Lemma 1. The proof relies on the Ostrowski's formula (1957). Define function $g = f^{-1}$. By Ostrowski's formula, we have

$$g^{(n)}(x) = \sum \frac{(-1)^j (n+j-1)!}{(f'(g(x)))^{n+j}} \prod_{i=2}^{n} \frac{1}{b_i!} \left(\frac{f^{(i)}(g(x))}{i!} \right)^{b_i}, \tag{20}$$

with $b_i \in \mathbb{N}$ and

$$\begin{aligned} b_2 + ... + b_n &= j \\ 2b_2 ... + nb_n &= n + j - 1. \end{aligned}$$

Manipulating expression (20) leads to

$$\begin{aligned} g^{(n)}(x) &= \sum \frac{(-1)^j}{(f'(g(x)))^{n+j}} (n+j-1)! \prod_{i=2}^{n} \frac{1}{b_i!} \left(\frac{(-1)^i f^{(i)}(g(x))}{i!} \right)^{b_i} \times (-1)^{ib_i} \\ &= \sum \frac{(-1)^{n-1}}{(f'(g(x)))^{n+j}} (n+j-1)! \prod_{i=2}^{n} \frac{1}{b_i!} \left(\frac{(-1)^i f^{(i)}(g(x))}{i!} \right)^{b_i}, \end{aligned}$$

as $\sum_{i=2}^{n} ib_i = n + j - 1$. Since for all integers $i \geq 2$, we have $(-1)^i f^{(i)}(g(x)) > 0$ and $f' > 0$, we easily conclude that for all $x \in \mathbb{R}_{++}$ and all $n \in \mathbb{N}^*$, $(-1)^{n+1} g^{(n)}(x) > 0$.

Appendix F

Proof of Proposition 7. The condition is necessary for $n = 3$. Then, note that from $(r(x))^B = (x - r(x))^A$, we obtain that for all $x \in \mathbb{R}_+$

$$r(x)(1 + r^{\alpha}(x)) = x, \text{ with } \alpha = \frac{B - A}{A} \in [0, 1].$$

Next, we want that show that $(-1)^n r^{(n)}(x) \leq 0$. Define auxiliary f with $f(x) = x(1 + x^{\alpha})$ and note that

$$\begin{aligned} f'(x) &= 1 + (1 + \alpha)x^{\alpha} > 0 \\ f''(x) &= (1 + \alpha)\alpha x^{\alpha - 1} > 0, \end{aligned}$$

and more generally for all integer $n \geq 2$ since $\alpha \in [0, 1]$, we have

$$(-1)^n f^{(n)}(x) > 0.$$

Using the result from lemma 1 leads to the desired result.

8 References

1. Abraham, A., S. Koehne amd N. Pavoni. (2011). On the fist-order approach in principal-agent models with hidden borrowing and lending. *Journal of Economic Theory* 146, 1331-1361

2. Alvi, E.(1997). First-order approach to principal—agent problems: A generalization. *The Geneva Papers on Risk and Insurance Theory* 22, 59-65.

3. Bernstein, S.N. (1928). Sur les fonctions absolument monotones. *Acta Mathematica* 52, 1-66.

4. Borch, K. (1962). Application of game theory to some problems in automobile insurance. *Astin Bulletin* 2(02), 208-221.

5. Caballé, J. and Pomansky, A. (1996). Mixed risk aversion, *Journal of Economy Theory* 71, 485-513.

6. Cato, S. (2013). The first-order approach to the principal-agent problems under inequality aversion. *Operations Research Letters* 41, 526-529.

7. Chaigneau, P., A. Edmans and D. Gottlieb. (2019). The informativeness principle without the first-order approach. *Games and Economic Behavior* 113, 743-755.

8. Conlon, John R. (2009). Two new conditions supporting the first-order approach to multisignal principal—agent problems. *Econometrica* 77(1), 249-278.

9. Diamond, P. and Stiglitz, J. E. (1974). Increases in risk and in risk aversion. *Journal of Economic Theory* 8(3), 337-360.

10. Ekern, S. (1980). Increasing Nth degree risk. *Economics Letters* 6(4), 329-333.

11. Eeckhoudt, L., Gollier, C. and Schneider, T. (1995). Risk aversion, prudence and temperance: A unified approach, *Economic Letters* 48, 331-336.

12. Faà di Bruno, F. (1857). Note sur une nouvelle formule de calcul differentiel (On a new formula of differential calculus), *The Quarterly Journal of Pure and Applied Mathematics* (in French) 1, 359-360.

13. Garicano, F. (2000). Hierarchies and the Organization of Knowledge in Production, *Journal of Political Economy*, 108(5), 874-904.

14. Hausdorff, F. (1921). Summationsmethoden und Momentfolgen I, *Mathematische Zeitschrift* 9, 74-109.

15. Hölmstrom, B. (1979). Moral hazard and observability. *The Bell Journal of Economics* 10, 74-91.

16. Hughes, J. S. (1982). Agency theory and stochastic dominance. *Journal of Financial and Quantitative Analysis* 17(03), 341-361.

17. Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3(4), 305-360.

18. Jewitt, I. (1988). Justifying the first-order approach to principal—agent problems. *Econometrica* 56(5), 1177-1190.

19. Jouini, E., Napp, C. and Nocetti. (2013). Economic Consequences of Nth Degree Risk Aversion Increase and Nth-Degree Risk Attitudes, Journal of Risk and Uncertainty 47, 199-224.

20. Jindapon, P. and Nielson, W. (2007). Higher-order generalizations of Arrow-Pratt and Ross risk aversion: A comparative statics approach. *Journal of Economic Theory* 136(1), 719-728.

21. Keenan, D. and Snow, A. (2002). Greater downside risk aversion, *Journal of Risk and Uncertainty* 24(3), 267-277.

22. Kimball, M. (1990). Precautionary saving in the small and in the large. *Econometrica* 58(1), 53-73.

23. Kimball, M. (1993). Standard risk aversion, Econometrica 61, 589-611.

24. LiCalzi, M. and Spaeter, S. (2003). Distributions for the first-order approach to principal-agent problems. *Economic Theory* 21(1), 167-173.

25. Liu, L., and Meyer, J. (2013). Substituing one risk increase for another: A method for measuring risk aversion. *Journal of Economic Theory* 148, 2706-2718.

26. Liu, L., and Neilson, W. (2019). Alternative approach to comparative nth-degree risk aversion. *Management Science*, https://doi.org/10.1287/mnsc.2018.3119.

27. Mas-Colell, A., Whinston, M. and Green, J. (1995). Microeconomic Theory, Oxford University Press.

28. Menezes, C., Geiss, C., and Tressler, J. (1980). Increasing downside risk. *The American Economic Review* 70(5), 921-932.

29. Menezes, C. F. and Wang, X. H. (2005). Increasing outer risk. *Journal of Mathematical Economics* 41(7), 875-886.

30. Mirrlees, J. A. (1976). The optimal structure of incentives and authority within an organization. *The Bell Journal of Economics* 7, 105-131.

31. Mirrlees, J. (1975). The theory of moral hazard and unobservable behaviour: Part I, Mimeo Nuffield College, Oxford. Published at the *Review of Economic Studies* (1996) 66, 3-21.

32. Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics* 5(2), 147-175.

33. Ostrowski, A. (1957). Le developpement de Taylor de la fonction inverse, *Comptes rendus de l'Académie des sciences Paris* 244, 429–430.

34. Pratt, J. (1964). Risk aversion in the small and in the large, *Econometrica* 32, 122-136.

35. Rogerson, W. (1985). The first-order approach to principal-agent problems. *Econometrica* 53(6), 1357-1367.

36. Ross, S.A., (1977). The Determination of Financial Structure: The Incentive-Signalling Approach. *Bell Journal of Economics* 8, 23-40.

37. Ross, S.A., (1981). Some stronger measures of risk aversion in the small and in the large with applications. *Econometrica* 49, 621-663.

38. Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory* 2(3), 225-243.

39. Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *Review of Economics Studies* 75, 957-984.

40. Sinclair-Desgagné, B. (1994). The first-order approach to multi-signal principal—agent problems. *Econometrica* 62(2), 459-465.

41. Widder, D., (1931). Necessary and sufficient conditions for the representation of a function as a Laplace integral. *Transactions of the American Mathematical Society* 33, 851-892.

42. Williamson, R., (1956). Multiply monotone functions and their Laplace Transform, *Duke Mathematical Journal* 23, 189-207.

43. Zeckhauser, R. (1970). Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory* 2(1), 10-26.