

Optimal "Pay-Performance-Sensitivity" in the presence of Exogenous Risk

Thomas Hemmer*

March, 2012

Abstract

I study the relation between the level of exogenous “risk” and pay-for-performance sensitivity (*PPS*) of optimal contracts. I first show that none of the known restrictions allow for exogenous risk in the standard principal-agent model. Next, I identify restrictions that support using the first-order approach for sufficiently high levels of exogenous risk. Third, by placing further restrictions on the agent’s preferences I provide a parametric example where there is not a monotone relation between risk and *PPS*. Finally, by appealing to the limiting case of this example I make the case that optimal contracts are unlikely to exhibit a monotone relation between risk and *PPS* even when risk is exogenous.

*Preliminary and incomplete. Please do not circulate. I appreciate comments and suggestions from Eva Labro, Christian Leutz, Doug Hanna, Bjorn Jorgensen as well as from workshop participants at Aarhus University, the FMG (London School of Economics), the University of Houston, OSU, SMU, UNC, the Utah Winter Conference and the Chicago/Minnesota Theory Conference. All errors (and I know there are some) are exclusively mine.

1 Introduction

Research on executive compensation in accounting and related fields relies to a significant extent on agency theory. Within this paradigm, the ability to study determinants and properties of optimal contracts rests on the ability to formulate the contracting problem in a tractable manner. For the standard principal-agent model that involves reliance on the so-called first order approach. Unfortunately, while the first-order approach offers the tractability needed, applying it comes at a cost. As is well known since (and due to) Mirrlees [1974], the approach is not generally valid and thus neither are the results the approach facilitates. As first proposed by Mirrlees, however, and subsequently proven by Rogerson [1985], validity of the first-order approach can be assured if attention is confined to production functions satisfying the so-called *MLR* and *CDF* conditions. Roughly, the former suggest that higher outcomes be more consistent with higher effort, while the latter corresponds to a form of stochastically diminishing return to effort in the production function.

While from a theoretical perspective these conditions have (some) intuitive appeal, from a more practical point of view these conditions are less satisfying as few examples of production functions that have these properties have been identified.¹ For example, the continuous equivalent of the discrete distribution with these properties provided in Rogerson [1985], while possessing the same desired properties, does not lend itself to the first-order approach as it is easily shown to suffer from the Mirrlees non-existence problem.² Jewitt [1988] takes a different approach to extending the reach of the first-order approach. By placing arguably somewhat weak additional restrictions on the contracting parties preferences, along with a slightly stronger condition on the likelihood ratio (concavely increasing) he provides access to a broader class of productions functions that satisfy (also) arguably somewhat less stringent conditions than the *CDF* condition.³ These distributions are all bounded from below and seem to provide reasonable representations of phenomena of significant interest in contract theory (and practice) such as (the empirical distribution of) stock prices. Despite these efforts to map out new territory agreeable by the first-order approach, however, it still

¹Recent noteworthy exceptions are the two groups of distributions introduced by LiCalzi and Spaeter [2003].

²Specifically, Stole [2001] suggests that a continuous version of the distribution of the distribution provided by Rogerson [1985], a form of Beta-distribution, satisfies the *MLR* and *CDF* conditions. While the particular distribution he suggests does not satisfy the *CDF* condition, it is easily verified that for Beta versions that do, f^a/f falls without limit as x approaches its lower limit.

³The *MLR* condition remains central though.

seems fair to suggest that the current territories occupied are small at best and quite constraining realistically.

Seemingly unfazed of the practical challenges involved in characterizing optimal incentive contracts in standard models of moral hazard, a large body of empirical work has emerged testing a central prediction attributed to the principal-agent theory: an inverse relation between “risk” or noisiness of a performance measure and “pay-for-performance sensitivity” (*PPS* hereafter). This relation, while often in the empirical literature illustrated using the contractual form of Holmstrom and Milgrom [1987], is typically ascribed to principal-agent theory in general without having encountered much of a theoretical challenge. In (fitting) contrast high profile empirical studies such as those by Jensen and Murphy [1990], Haubrich [1994], Garen [1994], Aggarwal and Samwick [1999], and Core and Guay [2002] have provided only limited or, perhaps more appropriately characterized, contradictory evidence on this relation.⁴

The insistence on an inverse equilibrium-relation between risk or noise and the sensitivity of pay to performance in optimal contracts targeted at the standard problem of moral hazard seems to be based on a view of risk being exogenous. More specifically, it appears based on implicitly restricting attention to production functions where the agent’s choice is confined to selecting only the mean of a symmetric, unimodal, continuous output distribution to allow for the exogenously determined dispersion to serve as a measure of “risk.” For such models the logical equilibrium link between risk and incentive then seems straight forward to establish. Presumably, if the dispersion of the outcome distribution increases, a (risk-neutral) principal would respond by decreasing the variation in the sharing rule to attain the optimal trade-off between improving risk-sharing and weakening of incentives.

While this link is intuitively appealing the simple fact is, however, that little is known about the nature of optimal contracts in production environments of this type. Indeed, due to methodological constraints, no results are available that link exogenous output risk to the *PPS* of optimal contracts for standard basic (one shot) principal-agent models. The purpose of this paper is to fill out methodological gaps that have stood in the way of doing this. Then, in turn, to use the methodological advances to achieve insights into the (equilibrium) “risk-*PPS*” trade-off predicted by the standard principal-agent model.

⁴See also Prendergast [2002] for an overview of this literature.

Several specific issues are addressed to this end. First, I establish that using the first-order approach to characterize the optimal risk/incentive trade-off for production functions of the “effort-plus-noise” type cannot be supported by neither the conditions provided by Mirrlees and Rogerson, nor by those provided by Jewitt [1988].⁵ Second, I then demonstrate that necessary conditions for the first-order approach to be valid for such production functions are that the likelihood ratio is bounded from below and that the dispersion of the noise-term (the risk) (therefore) is “sufficiently large.” This makes the comparative static exercise relating risk to the *PPS* of sharing rules derived using the first-order approach somewhat fragile. I do, however, demonstrate in a setting where the first-order approach *is* valid and *PPS* is well-defined, that the *PPS* of the optimal sharing rule is not monotonically decreasing in the dispersion of the noise term! Rather it is first increasing - then decreasing as risk is increased. Finally, suggestions as how to reexamine the mixed empirical results in the light of the findings provided here as well as some concluding remarks are offered.

2 Fixing Ideas

To help identify the key issues that must be addressed to formally link exogenous risk to *PPS*, consider Figures 1 through 3. Each of the three figures depicts the outcome density for a relatively high and low risk production environment where only the mean of each density is determined by effort, a . The figures differ only in terms of the level of effort supplied: the lowest possible, a_0 , in Figure 1, an intermediate level, a_m , in Figure 2 and in Figure 3 a higher level of effort, a^* , the significance of which I will return to a bit later. The (identical) kinked $u(s(x))$ -curve in the three figures illustrates outcome-dependent utility levels (ignoring cost of effort) under the optimal contract $s(x)$, where x is the (contractible) economic outcome.

Before proceeding, some remarks on the nature and the properties of the contract depicted here are in order. While the particular functional form with an “incentive zone” straddled by lower bound, \underline{U} , and a cap, \bar{U} , has some real world appeal, these specific features seem somewhat at odds with those of optimal contracts obtained from agency models. Accordingly, relying on this form here may appear somewhat arbitrary or even inappropriate. As will become apparent from the subsequent formal analysis, however, the key reason why such contracts are elusive in the agency

⁵Specifically I show that no production function of this type exists that satisfies either the Mirrlees/Rogerson or the Jewitt conditions for which the first-order approach is valid.

literature is exactly because the non-existence of model specifications that allow for analysis of the case where risk is truly exogenous. The specification partly introduced, partly derived in this paper contains as a naturally occurring case optimal contracts with the exact properties depicted in Figures 1 through 3. More generally, non-linearity as dictated by upper and lower bounds and some underlying symmetry around the second-best effort-level, a^* ,⁶ are the utility-space properties of all optimal contracts derived from the standard principal-agent model in cases where risk is truly exogenous.

Start comparing Figure 1 with Figure 2. The way Figure 1 is drawn here, the agent's expected utility is pretty much the lower bound for both of the risk levels since virtually all the probability mass resides below the bottom threshold of the incentive zone in either case. As effort is increased in Figure 2 and more probability mass is pushed onto the variable portion of the contract expected utility, Eu , increases in both cases. Notice however that for the level of effort depicted in Figure 2 the expected utility for the high risk distribution now is strictly higher than that for the low risk case as the upper bound is largely irrelevant for both distributions at that effort level. Indeed for levels of effort between a_0 and a_m , dEu/da for this particular contract is highest in the high risk case. In other words, the incentives provided by this contract in this effort-range are highest for the high-risk distribution.

Now compare instead Figures 2 and 3. Because in Figure 3 the two distributions now are centered on a^* half way between the start and end of the incentive zone, the expected utility is the same for both: $(\bar{u} - \underline{u})/2$. Accordingly, somewhere between a_m and a^* , dEu/da must be higher for the low-risk distribution than it is for the high-risk distribution. It should be clear, that this certainly is the case at a^* as the lower and upper bounds have a more significant (dampening) effect on dEu/da when the distribution is more disperse. Accordingly, in cases where risk is truly exogenous, the strengths of the incentives for a specific contract are not constant in effort. Moreover for any given level of effort the incentives provided by a given contract depend not just on the *PPS* of the contract itself - they also depend on the level of risk.

The implications of this for trying to link exogenous risk to *PPS* are two fold - one technical, one conceptual. From a technical perspective the problem is that the ability to identify the optimal contract generally requires reliance on the so-called first-order approach. To be able to exploit this

⁶Specifically, symmetry of the RHS of equation (7) in Holmstrom (1979).

approach it must be the case that the a for which $d(Eu - v(a))/da = 0$ is unique and is the one that is maximizing $Eu - v(a)$ over the contract one derives using the approach. Whether that is the case or not, however, will unfortunately itself hinge on the level of risk involved. To see this consider redrawing figures 1 through 3 for the case where the risk is approaching zero and all the probability mass therefore is mainly concentrated on a . In such a case a^* should be approaching the first-best and then be at least as high as in Figure 3. Further, as the distribution narrows, so does the incentive zone and virtually all the variation in the contract in this case will thus be concentrated in close proximity of a^* .

With these changes, consider moving a from a_0 to a^* . Because Eu remains basically constant as a is increased, $Eu - v(a)$ is actually decreasing between a_0 to a^* . The same is true for $a > a^*$, and since the contract by construction has $(d(Eu - v(a))/da)|_{a^*} = 0$, where $v(a)$ is the agent's convexly increasing cost of effort, a_0 and not a^* is actually the effort level that maximizes $Eu - v(a)$. Stated differently, in the limit as the risk goes to zero, for the contract obtained using the first-order approach the agent's expected utility at a^* is at a saddle-point with Eu declining both before and after a^* . Consequently the approach used to characterize the contract is invalid. If the distributions (and thus the incentive zone) in figure 1 are instead widened, however, it may then well be the case that Eu is everywhere increasing between a_0 to a^* in which case the *FOA* could well be valid. Accordingly, before basing any conclusions about the relation between truly exogenous risk and *PPS* on the *FOA*, the validity of the *FOA* must first be ensured.

Conceptually what is important here is that incentive strength and *PPS* cannot be equated even when the *FOA* is valid exactly due to the non-linearities introduced by the required boundedness of the likelihood ratio. Increasing the variance of either of the distributions in Figure 1 reduces $(d(Eu - v(a))/da)|_{a^*}$ even if no changes are made to the contract itself. Assume for specificity then that a^* is the second-best effort level for the low variance distribution. Clearly, then in the high risk case the optimal response to this contract is to choose an a lower than a^* . Shifting the contract to the left would improve the incentives and then move the high-risk a closer to a^* but the highest a that can be implemented in the high risk case with this level of *PPS* will be strictly lower than a^* . The problem then is of-course that if the highest a this contract can implement when optimally centered is less than high-risk second-best, it implies that it could be optimal to increase *PPS* when risk is increased. As I'll show in this paper this is exactly the case for a class of production

functions with exogenous risk for which the *FOA* is actually valid.

3 The Basic Model

Given the issues at hand are optimal sharing rules for production functions where risk is entirely exogenous and, in particular, the ability of the first-order approach to deliver such, the production functions to be considered here are on the form

$$x = a + \epsilon(\theta), \tag{1}$$

where $a \in \mathcal{R}^+$ is the agent's "effort," and $\epsilon(\theta)$ is the random (or risky) component parameterized by θ - the determinant of the dispersion of the random component. Here, lower a value of θ is taken to imply that $\epsilon(\theta)$ is less risky in the sense of second-order stochastic dominance. I use $f(x, a, \theta)$ to denote the density of x for given values of a and θ while $F(x, a, \theta)$ is used to represent the corresponding distribution function. $F(x, a, \theta)$ and $f(x, a, \theta)$ are as usual both required here to be twice continuously differentiable in a .

For simplicity and consistent with the literature relating risk to *PPS*, the principal is taken to be risk-neutral while the risk-averse and effort-averse agent has (standard) additively separable preferences, $U(s(x), a) \equiv u(s(x)) - v(a)$ where $s(x)$ is the agent's share of realized output, $u(s(x))$ is the agent's increasing and strictly concave utility for income defined over the entire real line, and $v(a)$ is the agent's strictly increasing, convex, and twice continuously differentiable cost of effort. $h(u(s(x))) \equiv s(x)$ is the inverse of the agent's utility of consumption. The agent's reservation utility is represented by \underline{U} and the timing of events is the standard one.

For an agency of this form the following observation (adapted from Jewitt [1988]) is of central importance:

Observation 1. *When the production function is of the form (1), then $F(x, a, \theta) = F(x - a, \theta)$.*

Indeed, this observation forms the basis for the first result:

Proposition 1. *No production function of the form (1) satisfying either (i) the Jewitt-conditions or (ii) the Mirrlees-Rogerson-conditions exists for which the first-order approach (FOA) to characterizing the optimal solution to the principal's problem is valid.*

Proof of Proposition 1. First, note that when $f(x, a, \theta) = f(x - a, \theta)$, the support of x must be unbounded from below to avoid moving support and, thus, first-best. To prove part (i) then, suppose that the *FOA* is valid so that the optimal contract satisfies

$$\frac{1}{u'(s(x))} = \lambda + \mu \frac{f^a(x, a, \theta)}{f(x, a, \theta)}, \quad (2)$$

where λ and μ are the Lagrange multipliers associated with the (*IR*) and the (*IC*) constraints respectively. With the support of x unbounded from below, condition (2.11) of Jewitt [1988] implies that $f^a(x, a, \theta)/f(x, a, \theta) \rightarrow -\infty$ as $x \rightarrow -\infty$. Since validity of the *FOA* implies that $\mu > 0$,⁷ the RHS of (2) falls without limit as $x \rightarrow -\infty$. But the LHS of (2) is always strictly positive contradicting the presumption that the *FOA* is valid. This concludes the first part of the proof.

To prove part (ii) note that $F^{aa}(x - a) = f^x(x - a)$. *CDFC* thus requires $f^x(x - a)$ to be everywhere non-negative and strictly positive for some x .⁸ Since the support of x must be unbounded from below and since $\int_{-\infty}^{\bar{\epsilon}+a} f(x - a)dx = 1$, where $\bar{\epsilon}$ is the upper bound of the support of $f(x)$, *CDFC* implies that $\bar{\epsilon}$ must be finite and that $f(\bar{\epsilon}) > 0$. Accordingly, then $f(x - a)$ is not differentiable w.r.t. a at $x = \bar{\epsilon} + a$. *Q.E.D.*

4 Alternative Restrictions

Given the incompatibility of production functions of the form (1) with the standard restrictions used to ensure the validity of the *FOA*, any insight into the relation between exogenous risk and *PPS* requires identification of alternate properties to be possessed by the production function and the agent's preferences. The goal, is of course to ensure that the a^* that solves $d(Eu - v(a))/da = 0$ is unique and that $(d^2(Eu - v(a))/da^2)|_{a^*} < 0$. At the outset it seems quite intuitive that this would be hard to guarantee here unless the resulting optimal sharing rule is monotone non-decreasing in x . As a starting point it therefore seems natural to confine attention to production functions that satisfy the *MLR* condition.⁹ This immediately translates into what is arguably a desirable property

⁷Holmstrom (1979) Proposition 1. Also, Jewitt (1988) established *without* first assuming that the *FOA* is valid that in the case of a risk-neutral principal, $\mu > 0$ if a is interior.

⁸Jewitt (1988).

⁹It follows directly from (1) that when the first-order approach is valid, $s(x)$ is monotone non-decreasing in x if and only if f^a/f is monotone non-decreasing in x .

of $f(x, a, \theta)$:

Lemma 1. *When the production function is of the form (1), the MLR condition is satisfied if and only if $f(x, a, \theta)$ is unimodal.*

Proof of Lemma 1 Suppose $f(x, a, \theta)$ is not unimodal. Then, for any given a and θ there exists corresponding values of x , say \underline{x} and \bar{x} , $\underline{x} < \bar{x}$, such that $f^x(x, a, \theta)|_{x=\underline{x}} < 0$ and $f^x(x, a, \theta)|_{x=\bar{x}} > 0$. Since $f^x(x, a, \theta) = -f^a(x, a, \theta)$ when the production function is of the form (1), $f^a(\underline{x}, a, \theta)/f(x, a, \theta) > f^a(\bar{x}, a, \theta)/f(x, a, \theta)$. *Q.E.D.*

For the purpose of concreteness and to facilitate further analysis I confine my attention to a class of unimodal distributions that are also symmetric. Accordingly, the production functions under consideration satisfy:

$$f(x, a, \theta) = g\left(\frac{-b(|x-a|)}{\theta}\right) k(\theta), \quad (3)$$

where $g' > 0$, $g(0) > 0$ and finite, $\lim_{\frac{b(|x-a|)}{\theta} \rightarrow \infty} g\left(\frac{-b(|x-a|)}{\theta}\right) \rightarrow 0$, $b(0) = 0$, $b'(0) = 0$, $b' > 0 \forall x \neq a$, and $k(\theta) > 0$ satisfies $d \int_{-\infty}^{\infty} g\left(\frac{-b(|x-a|)}{\theta}\right) k(\theta) dx / d\theta = 0$. Besides arguably being a natural choice, confining attention to symmetric distributions is motivated by two factors. First, comparing the riskiness of skewed distributions is more involved and places restrictions on the type of skewness that can be considered anyway. Second, the empirical work on the relation between risk and *PPS* typically make reference to the variance of a symmetric distribution as the proper measure of risk.

That the likelihood ratio must be monotone and, from the proof of Proposition 1, is required to be bounded then places further restrictions on b and g :

Lemma 2. *For the class of production functions given by (3), the likelihood ratio is monotonic non-decreasing and bounded if and only if (i) $b'g'/g$ is monotonic non-decreasing in x and (ii) $f(x, a, \theta) \rightarrow k(\theta) \text{Exp}[-B|x-a|/\theta + R]$ as $|x-a|/\theta \rightarrow \infty$.*

Proof of Lemma 2. With (3) we have

$$f^a = \text{Sign}[x-a] \frac{b'}{\theta} g' k(\theta),$$

so that

$$\frac{f^a}{f} = \text{Sign}[x-a] \frac{b' g'}{\theta g}, \quad (4)$$

which is monotonic non-decreasing in x if and only if $b'g'/g$ is. Moreover, given (4) $\lim_{(|x-a|/\theta) \rightarrow \infty} \frac{f^a}{f} = B > 0$ implies $\lim_{(|x-a|/\theta) \rightarrow \infty} \frac{g'}{g} = \frac{B}{b'}$ or that $\frac{1}{g}dg = Bd(|x-a|/\theta)$. Solving yields the second requirement in Lemma 2. *Q.E.D.*

Corollary 1. *For the class of production functions given by (3) with likelihood ratios bounded above at $B > 0$, $\lim_{(|x-a|/\theta) \rightarrow \infty} b'(\cdot) \rightarrow B$.*

Lemma 3 *As $\theta \rightarrow 0$,*

$$\frac{d(f^a/f)}{dx} \rightarrow \begin{cases} 0, & \text{for } x \neq a, \\ \infty, & \text{for } x = a. \end{cases}$$

Proof of Lemma 3 Since $\frac{d(f^a/f)}{dx} = \frac{b''g'}{\theta g} + \frac{b'}{\theta} \left[\frac{g''}{g} - \left(\frac{g'}{g} \right)^2 \right]$, Lemma 2 yields the first line. With $\frac{d(f^a/f)}{dx}|_{(|x-a|/\theta)=0} = \frac{b''g'}{\theta g}|_{(|x-a|/\theta)=0}$ the second line follows from the fact that with $f(x, a, \theta)$ twice continuously differentiable in x , $\frac{b''g'}{g}|_{(|x-a|/\theta)=0} > 0$ and finite. *Q.E.D.*

Unfortunately, the necessary restrictions identified above rule out many (if not all) of the best known standard distributions that could be useful as a platform for further analysis. To identify distributions that do belong to the class of candidate distributions defined by (3) and Lemma 2 consider first the (Double Exponential) Laplace distribution:

$$f(x, a, \theta) = \frac{1}{2\theta} e^{-\frac{|x-a|}{\theta}}, \quad (5)$$

which satisfies part (ii) of Lemma 2.

Although this distribution perhaps is somewhat unusual from an empirical vantage point, the specific property of (5) that f^a/f is a negative (positive) constant for $x < (>) a$, is quite useful for identifying the general fallacy in asserting that the level of exogenous risk must be inverse related to the *PPS* in optimal incentive contracts.¹⁰ Unfortunately (5) doesn't satisfy the requirement that $b'(0) = 0$ and $f(x, a, \theta)$ is thus not twice differentiable in a (or x) at $a = x$. Fortunately, however, it is straight forward to obtain a density function that is everywhere twice continuously differentiable in a (and x) that preserves the properties of the Laplace distribution almost everywhere.

To this end, consider the following hybrid distribution:

$$\gamma f(x, a) = \begin{cases} N(x, a) = n + \frac{\psi}{\omega\theta} e^{-\frac{\eta(x-a)^2}{2\theta}}, & a - \omega \leq x \leq a + \omega, \quad \omega > 0, \\ L(x, a) = \frac{\eta}{\theta} e^{-\frac{|x-a|}{\theta}}, & \text{otherwise,} \end{cases} \quad (6)$$

¹⁰I will return to this in the Discussion section.

where $\gamma \equiv 2 \left(\int_a^\omega N(x, a) dx + \int_\omega^\infty L(x, a) dx \right)$. Both $N(x, a)$ and $L(x, a)$ are twice continuously differentiable at $x = a \pm \omega$ and symmetric around a while $N(x, a)$ also is twice continuously differentiable at $x = a$. For a production function of the form (6) to be everywhere twice continuously differentiable in a it is then required that

$$N(\omega) = L(\omega),$$

$$\left. \frac{dN(x-a)}{da} \right|_\omega = \left. \frac{dL(x-a)}{da} \right|_\omega,$$

and

$$\left. \frac{d^2 N(x-a)}{da^2} \right|_\omega = \left. \frac{d^2 L(x-a)}{da^2} \right|_\omega.$$

Solving yields $\eta = \frac{1+\theta/\omega}{\omega}$, $n = \frac{1}{\omega^2} e^{-\frac{\omega}{\theta}}$ and $\psi = e^{-\frac{\omega-\theta}{2\theta}}$. With this I then have

$$N(x, a) = \frac{1}{\omega\theta} e^{-\frac{(1+\theta/\omega)(x-a)^2 + \omega^2 - \omega\theta}{2\theta\omega}} + \frac{1}{\omega^2} e^{-\frac{\omega}{\theta}}.$$

$$\frac{f^a(x, a)}{f(x, a)} = \begin{cases} \frac{(1/\omega\theta + 1/\omega^2)(x-a)}{1 + \frac{\theta}{\omega} e^{\frac{(1/\omega + \theta/\omega^2)(x-a)^2 - \omega - \theta}{2\theta}}}, & a - \omega \leq x \leq a + \omega, \\ \frac{\text{Sign}[x-a]}{\theta}, & \text{otherwise.} \end{cases}$$

Notice now that *if* the *FOA* is valid when the production function is on the form (6), as $\omega \rightarrow 0$, the optimal contract approaches

$$u(s(x)) = \begin{cases} \underline{U} + v(a^\theta) + \theta v'(a^\theta), & \text{for } x \geq a^\theta, \\ \underline{U} + v(a^\theta) - \theta v'(a^\theta), & \text{for } x < a^\theta, \end{cases} \quad (7)$$

independent of the relation between u and $1/u!$ ¹¹ The lack of dependence between $s(x)$ and the properties of u follows since as when $\omega \rightarrow 0$, $[\lambda + \mu \frac{f^a}{f}]$ basically takes on only one of two different values while both the (*IR*) and the (*IC*) constraint still must be (exactly) satisfied. To exploit this simplicity, in what follows I focus on the case where $f(x, a)$ is on the form (6) with ω being arbitrarily close to zero (and thus arbitrarily close to the production function given by (5)) and where the contract derived via the *FOA* therefore is arbitrarily closely approximated by the $s(x)$ that satisfies (7). In what follows I will take the (approximate) properties of (5) and (7) to be exact

¹¹ a^θ denotes the agent's optimal response to the optimal contract (derived using the *FOA*) for a given θ .

and will refer to this combination as the Approximate Laplace Specification.

While a main attraction of this specification is the simplicity of the contract derived from the *FOA*, equally important for the purpose at hand is that the risk-*PPS* trade-off when the *FOA* is valid becomes well defined. Specifically, with this representation the *PPS* is measured unambiguously by the difference in the payments for $x < a^\theta$ and $x \geq a^\theta$. The standard empirical hypothesis that increased risk here measured by the Laplace distribution's standard deviation, θ , should be accompanied by a decrease in the *PPS* in the agent's contract, therefore simply becomes a statement about the equilibrium relation between θ and $h(\underline{U} + v(a^\theta) + \theta v'(a^\theta)) - h(\underline{U} + v(a^\theta) - \theta v'(a^\theta))$. The remainder of this section explores this relation and in doing so demonstrates the lack of a monotone relation between risk and *PPS*.

Since the *PPS* in (7) is monotonic increasing in:

$$\theta v'(a^\theta), \tag{8}$$

it would seem to be straight forward to conclude that the relation between exogenous risk and *PPS* cannot be everywhere decreasing. This appears to follow because when $\theta \rightarrow 0$, presumably $a \rightarrow a^{FB}$ and with $v'(a^{FB}) \leq 1$ for the production function specified here, $\theta v'(a^\theta) \rightarrow 0$ from above as $\theta \rightarrow 0$.¹²

The problem with the above argument is closely related to Lemma 2. As $\theta \rightarrow 0$, $\partial EU/\partial a \rightarrow -v'(a)$ for any $a < a^{FB}$. Moreover, as $s(x) \rightarrow \underline{U} + v(a^{FB})$, $EU(s(x), 0) \rightarrow \underline{U} + v(a^{FB})$ as well. As under the optimal contract $EU(s(x), a^\theta) = \underline{U}$ independent of θ , in the limiting case of $\theta = 0$, $a = 0$ then is the agent's optimal response to the contract (7) implying that the first-order approach is not valid in this case. Accordingly, establishing that the relation between risk and *PPS* is not always everywhere decreasing by appealing to (8) requires that it be established that preferences exist such that the first-order approach is valid for a range of values of θ "sufficiently close to 0." Proposition 2 first addresses the possibility that the *FOA* is valid for any given value of $\theta > 0$.

Proposition 2. *For the Approximate Laplace Specification the FOA is valid at θ if $v(a)$ satisfies $v'(a^\theta)e^{\frac{a-a^\theta}{\theta}} - v'(a) > 0$, $a < a^\theta$.*

Proof of Proposition 2. A sufficient condition for the first-order approach to be valid in the case

¹² a^{FB} is here used to signify the first-best effort level.

of the Approximate Laplace Specification is that $\partial EU(s(x), a)/\partial a > 0$ for all $a < a^\theta$, where $s(x)$ is derived using the first-order approach. This follows since by construction, $\partial EU(s(x), a)/\partial a = 0$ for $a = a^\theta$ and, as is easily verified, for the Approximate Laplace Specification, $\partial EU(s(x), a)/\partial a < 0$ for $a > a^\theta$. With the structure of this contract,

$$EU(s(x), a) = \underline{U} + v(a^\theta) - \int_{-\infty}^{a^\theta} \theta v'(a^\theta) f(x, a, \theta) dx + \int_{a^\theta}^{\infty} \theta v'(a^\theta) f(x, a, \theta) dx - v(a),$$

and

$$\partial EU(s(x), a)/\partial a = 2\theta v'(a^\theta) f(a^\theta, a, \theta) - v'(a).$$

Accordingly, the first-order approach is valid if $v(a)$ satisfies

$$v'(a) < 2\theta v'(a^\theta) f(a^\theta, a, \theta), \quad a < a^\theta,$$

or, using (5),

$$v'(a^\theta) e^{\frac{a-a^\theta}{\theta}} - v'(a) > 0, \quad a < a^\theta.$$

As $v'(a^\theta) > 0$, $e^{\frac{a-a^\theta}{\theta}} > 0$, and $e^{\frac{a-a^\theta}{\theta}}$ is increasing convexly in a for $a < a^\theta$, it is always possible to find an increasing and convex cost function, $v(a)$, that satisfies this condition. *Q.E.D.*

Proposition 2 simply establishes that for the Approximate Laplace Specification, given a $\theta > 0$ and a solution a^θ satisfying the agent's first-order condition in a , this is indeed the optimal solution to the principal's problem as long as the agent's cost function does not increase "too quickly" in a for $a < a^\theta$. Proposition 3 establishes the implications of this constraint for the relation between θ and a^θ .

Proposition 3. *Suppose for some θ , $v(a)$ is such that $v'(a^\theta) e^{\frac{a-a^\theta}{\theta}} - v'(a) > 0$, $a < a^\theta$. Then for the Approximate Laplace Specification, $\frac{da^\theta}{d\theta} < 0$.*

Proof of Proposition 3. Given the nature of (7), the principal's residual maximization problem can be expressed as

$$\max_a \quad a - \frac{1}{2}h[\underline{U} + v(a) - \theta v'(a)] - \frac{1}{2}h[\underline{U} + v(a) + \theta v'(a)]$$

which has the first-order condition (suppressing (a) from the notation for the moment)

$$FOC \equiv 2 - (v' - \theta v'')h'[\underline{U} + v - \theta v'] - (v' + \theta v'')h'[\underline{U} + v + \theta v'] = 0.$$

Furthermore, taking the partial derivatives of the left-hand-side of this first-order condition with respect to θ and a respectively yield

$$\frac{\partial FOC}{\partial \theta} = v''h'[L] + v'(v' - \theta v'')h''[L] - v''h'[H] - v'(v' + \theta v'')h''[H], \quad (9)$$

and

$$\frac{\partial FOC}{\partial a} = -(v'' - \theta v''')h'[L] - (v' - \theta v'')^2h''[L] - (v'' + \theta v''')h'[H] - (v' + \theta v'')^2h''[H], \quad (10)$$

where $L \equiv \underline{U} + v(a^\theta) - \theta v'(a^\theta)$, and $H \equiv \underline{U} + v(a^\theta) + \theta v'(a^\theta)$. Accordingly,

$$\frac{da^\theta}{d\theta} = -\frac{v''h'[L] + v'(v' - \theta v'')h''[L] - v''h'[H] - v'(v' + \theta v'')h''[H]}{-(v'' - \theta v''')h'[L] - (v' - \theta v'')^2h''[L] - (v'' + \theta v''')h'[H] - (v' + \theta v'')^2h''[H]}.$$

where the denominator is the second-order condition for a^θ and thus guaranteed to be negative and the sign of $\frac{da^\theta}{d\theta}$ is therefore determined by the sign of the numerator.

The requirement that $v'(a^\theta)e^{\frac{a-a^\theta}{\theta}} - v'(a) > 0$, $a < a^\theta$ can be re-written as $v'(a^\theta)e^{\frac{a-a^\theta}{\theta}} - \delta(a, \theta) = v'(a)$, where $\delta(a, \theta) > (=) 0$ for $a < (=) a^\theta$. Accordingly, $\delta^a(a, \theta) \leq 0$, $a < a^\theta$. Now differentiate $v'(a)$ to get $\theta v''(a) = v'(a^\theta)e^{\frac{a-a^\theta}{\theta}} - \theta \delta^a(a, \theta)$. Accordingly, the requirement that $v'(a^\theta)e^{\frac{a-a^\theta}{\theta}} - v'(a) > 0$ implies that $\theta v''(a^\theta) \geq v'(a^\theta)$. This confirms that the numerator of the expression for $\frac{da^\theta}{d\theta}$ is also negative. *Q.E.D.*

Using the relation between a and θ from proposition 3, proposition 4 expands on proposition 2 to establish that *if* the agent's cost function supports using the FOA for a given $\hat{\theta} > 0$, the FOA is also valid for all $\theta > \hat{\theta}$.

Proposition 4. *For any Approximate Laplace Specification where $v(a)$ satisfies $v'(a^\theta)e^{\frac{a-a^\theta}{\theta}} - v'(a) > 0$, $a < a^\theta$, where a^θ satisfies the agent's first-order condition in a , the FOA is valid for all $\theta \geq \hat{\theta}$.*

Proof of Proposition 4. Consider some strictly positive value of θ , say $\hat{\theta}$, and define θ_ε as the

value of θ for which $a^{\theta_\varepsilon} = a^{\hat{\theta}} - \varepsilon$, $\varepsilon \geq 0$. Then, if a^θ is decreasing in θ , $\theta_\varepsilon \geq \hat{\theta}$. I can then re-write the condition from proposition 2 as

$$e^{\frac{-|a^{\hat{\theta}} - \varepsilon - a|}{\theta_\varepsilon}} - \frac{v'(a)}{v'(a^{\theta_\varepsilon})} > 0, \quad a < a^{\theta_\varepsilon} \text{ and } \varepsilon = 0. \quad (11)$$

Now differentiate the LHS of (11) with respect to ε to obtain

$$\left(\frac{1}{\theta_\varepsilon} + \frac{d\theta_\varepsilon}{d\varepsilon} \frac{(a^{\hat{\theta}} - \varepsilon - a)}{\theta_\varepsilon^2} \right) e^{\frac{-|a^{\hat{\theta}} - \varepsilon - a|}{\theta_\varepsilon}} + \frac{v'(a)v''(a^{\theta_\varepsilon})}{(v'(a^{\theta_\varepsilon}))^2} > 0, \quad \forall a \leq a^{\hat{\theta}} - \varepsilon > 0.$$

Accordingly, if the first-order approach is valid for $\theta = \hat{\theta}$, it is valid for all $\theta \in [\hat{\theta}, \infty)$. *Q.E.D.*

5 The “Risk - *PPS*” Trade-off

While the preceding analysis establishes that it is indeed feasible to place specific restrictions on preferences and the production function to guarantee the validity of the *FOA*, the task of establishing generally a relation between risk and *PPS* does not appear within reach, even with the benefits offered by the Approximate Laplace Specification. The reason is that the validity of the *FOA* is crucially dependent on the risk in the production function in that for any given set of preferences the validity of the *FOA* requires that the risk is not “too low.” In this section I therefore restrict the functional forms further to facilitate a specific (counter-) example to demonstrate that the maintained proposition that optimal contracts generally should exhibit an inverse relation between *PPS* and exogenous risk is false.

To that end I continue to rely on the Approximate Laplace Specification but utilize the following specific representation of the agent’s preferences:

$$U(s(x), a) = (s(x))^{\frac{1}{2}} - a^2/2,$$

and set $\underline{U} = 1$. With this representation, under the maintained assumption that the *FOA* is valid, the optimal level of effort, a^θ , is simply the solution to the following polynomial:

$$1 - 2a - a^3 - 2\theta^2 a = 0.$$

Table 1 below summarizes the solution along with the properties of the contract derived using the *FOA* as well as indicates the sign of the second-order condition (*SOC*) and accordingly the validity of the *FOA* at the various levels of risk as per the condition in proposition 2.

θ	a^θ	$2\theta v'(a^\theta)$	$s(H) - s(L)$	<i>SOC</i>
.1	0.44995	0.08999	0.19820	+
.2	0.43986	0.17594	0.38593	+
.3	0.42380	0.25428	0.55423	+
.4	0.40285	0.32228	0.69686	+
.5	0.37834	0.37834	0.81084	-
.6	0.35166	0.42199	0.89617	-
.7	0.32414	0.45380	0.95527	-
.8	0.29690	0.47504	0.99196	-
.9	0.27076	0.48737	1.01047	-
1.0	0.24627	0.49254	1.01495	-
1.1	0.22371	0.49216	1.00896	-
1.2	0.20230	0.48552	0.99091	-
1.3	0.18470	0.48022	0.97682	-
1.4	0.16812	0.47074	0.95477	-
1.5	0.15329	0.45987	0.93055	-
1.6	0.14006	0.44819	0.90518	-
1.7	0.12826	0.43608	0.87934	-
1.8	0.11773	0.42383	0.85353	-
1.9	0.10832	0.41162	0.82806	-
2.0	0.09990	0.39960	0.80319	-

The column to the right indicates if the *FOA* is (-) or is not (+) valid for the given level of θ . The third and fourth columns of this table provides the *PPS* in the contract respectively in terms of utiles and consumption units. The latter's relation with risk as measured by θ is plotted

in Figure 4 below.

Insert Figure 4 about here.

The solid curve in this figure represents the relation between risk and PPS over the range of θ where the FOA is indeed valid and the PPS listed in the table therefore is the PPS of the optimal contract. As reflected in this figure, the validity of the FOA hinges on the value of θ . For the particular preference representation the FOA attains validity for a value of θ between .4 and .5 and is, as per proposition 4, (in-) valid for all (smaller) greater values of θ .

What is most important about this example is, however, that it establishes that the PPS is not monotonically decreasing in θ over the range where the FOA is actually valid. This is only the case for relatively large values of θ while the relation between risk and PPS is actually reversed for relatively small values of θ . Unless the level of risk somehow is expected to always be “large” there is no theoretical reason for expecting that higher risk should be statistically associated with lower PPS . Indeed, the theory predicts mixed and/or weak empirical results unless, perhaps, the sample is partitioned on risk.

6 Discussion and Conclusion

The analysis in this paper provides several insights. First, it is possible to place restrictions on preferences and production functions in such a way that the optimal sharing rule can be derived and studied for the standard principal agent model with exogenous risk. Second, there is no theoretical support for the claim that the standard principal-agent model predicts an inverse relation between exogenous risk and PPS . In the specific case studied, it was shown that when risk is sufficiently low, a positive relation exists. Third, and somewhat depressing, finding general restrictions to validate this type of analysis is unlikely. This follows since the restrictions needed on the agent’s preferences depend on the nature of the optimal contract and except in the case of the Approximate Laplace Specification where the functional form of the optimal contract is extremely simple, identifying even sufficient restrictions is not likely to be doable.

Despite the lack of restrictions that would allow for a more general analysis of optimal contracts in settings with exogenous risk, however, the nature of the optimal relation between PPS and such exogenous risk is bound to be the same as in the special case identified and analyzed here. This

follows from the fact that any viable production function must satisfy lemma 2, implying that the optimal contracts are bounded from above as well as from below. As in the limit when the risk goes to zero, all contracts and all distributions “look” the same, the *PPS* must be approaching zero as in the case presented here. Similarly, as risk grows without limit no effort will be elicited and the optimal *PPS* then should be approaching 0 as well. Thus, while whether the optimal contracts can actually be derived in alternative settings remains an open question, the analysis here strongly suggests that if settings exist where this is feasible, the contracts will also exhibit a non-monotonic relation between risk and *PPS*.

References

- [1] Aggarwal, R., and A. Samwick, 1999, The Other Side of the Trade-Off: The Impact of Risk on Executive Compensation. *The Journal of Political Economy* 107, 65-105.
- [2] Core, J., and W. Guay, 2002, The Other Side of the Trade-Off: The Impact of Risk on Executive Compensation: a Comment. Working Paper, Wharton.
- [3] Garen, J., 1994, Executive Compensation and Principal-Agent Theory. *The Journal of Political Economy* 102, 1175-1190.
- [4] Haubrich, J., 1994, Risk Aversion, Performance Pay, and the Principal-Agent Problem. *The Journal of Political Economy* 102, 258-276.
- [5] Holmström, B., 1979, Moral Hazard and Observability. *Bell Journal of Economics* 10, 74-91.
- [6] Holmström, B., and P. Milgrom, 1987, Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica* 55, 303-28.
- [7] Jensen, M., and K. Murphy, 1990, Performance Pay and Top Management Incentives. *The Journal of Political Economy* 98, 225-264.
- [8] Jewitt, I., 1988, Justifying the First-Order Approach to Principal-Agent Problems. *Econometrica* 56, 1177-1190.
- [9] LiCalzi, M. and S. Spaeter, 2003, Distributions for the first-order approach to principal-agent problems. *Economic Theory*, Springer, vol. 21(1), 167-173.
- [10] Mirrlees, J., 1974, Notes on Welfare Economics, Information and Uncertainty. In Balch, McFadden, and Wu, editors, *Essays on Economic Behavior under Uncertainty*, Amsterdam: North Holland.
- [11] Prendergast, C., 2002, The Tenuous Trade-Off of Risk and Incentives. *The Journal of Political Economy*.
- [12] Rogerson, W., 1985, The First-Order Approach to Principal-Agent Problems. *Econometrica* 53, 1357-1367.
- [13] Stole, L., 2001, *Lectures on the Theory of Contracts and Organizations*, Working Paper, Chicago.

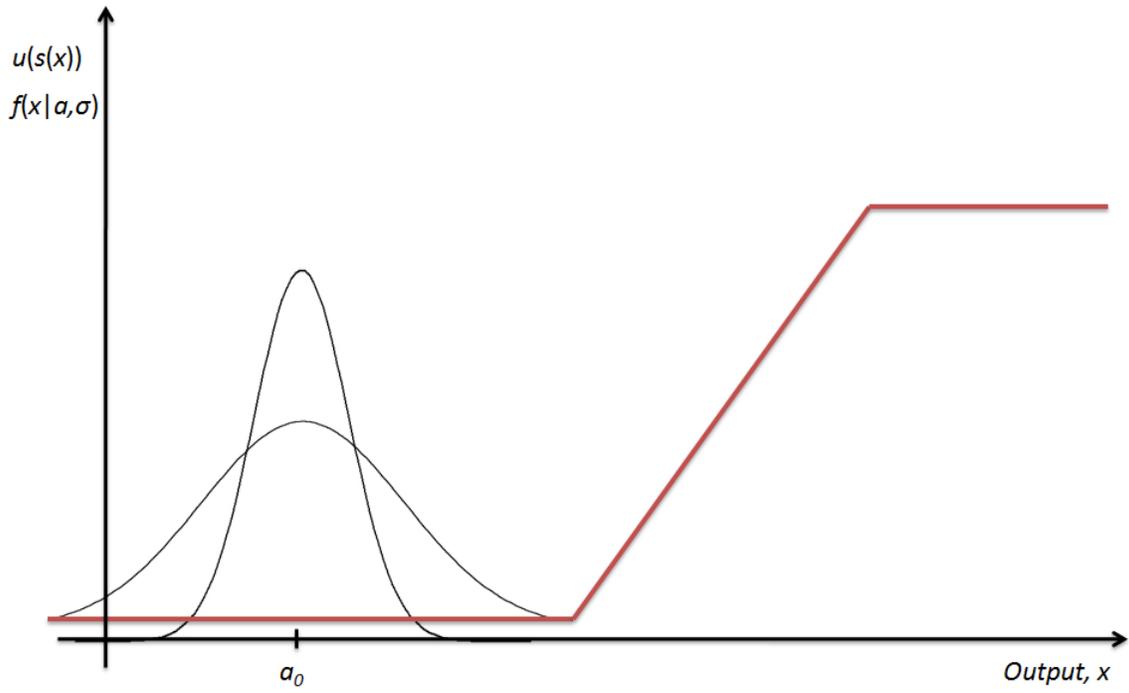


Figure 1: Figure 1: Minimum Effort

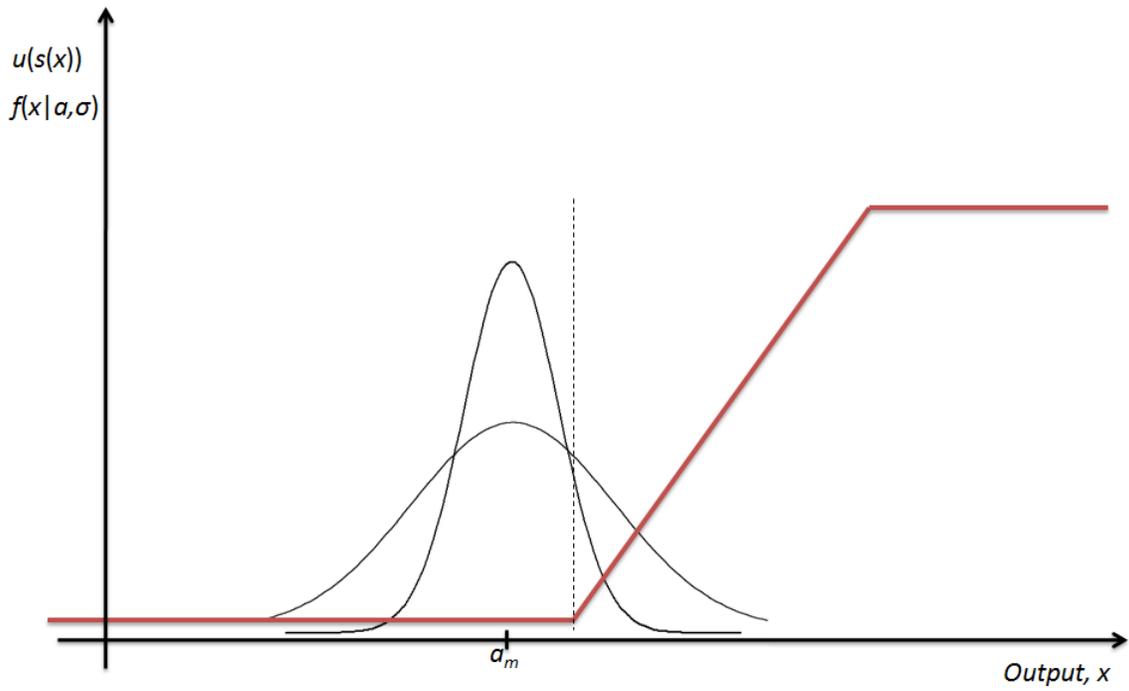


Figure 2: Figure 2: Medium Effort

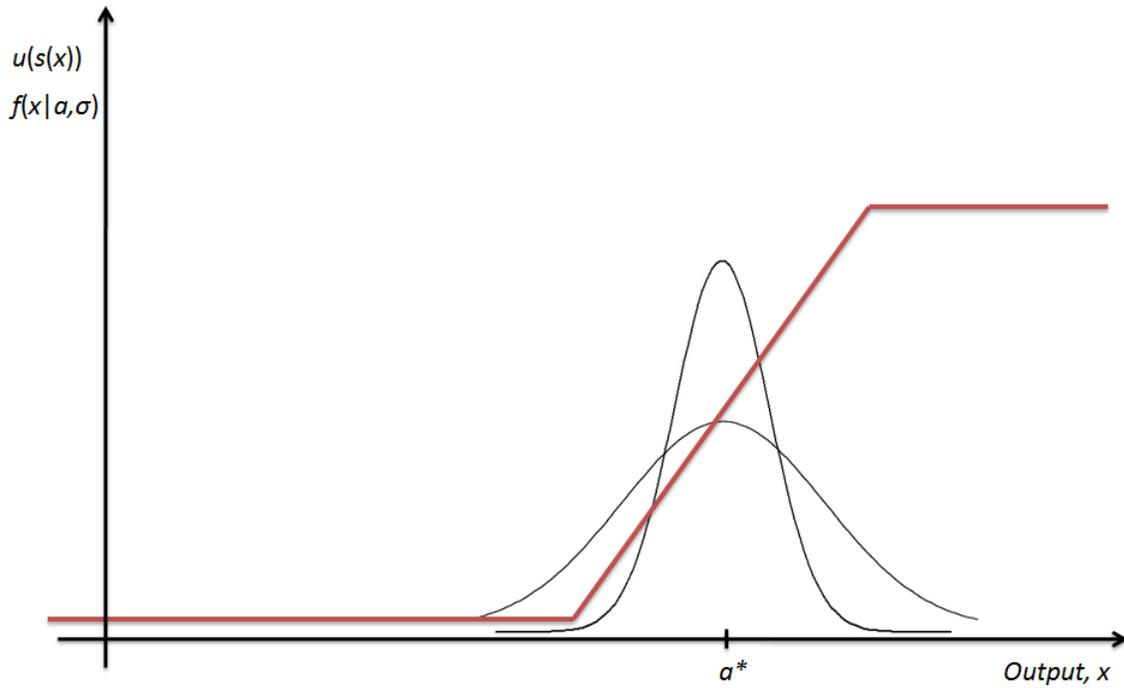


Figure 3: Figure 3: High effort

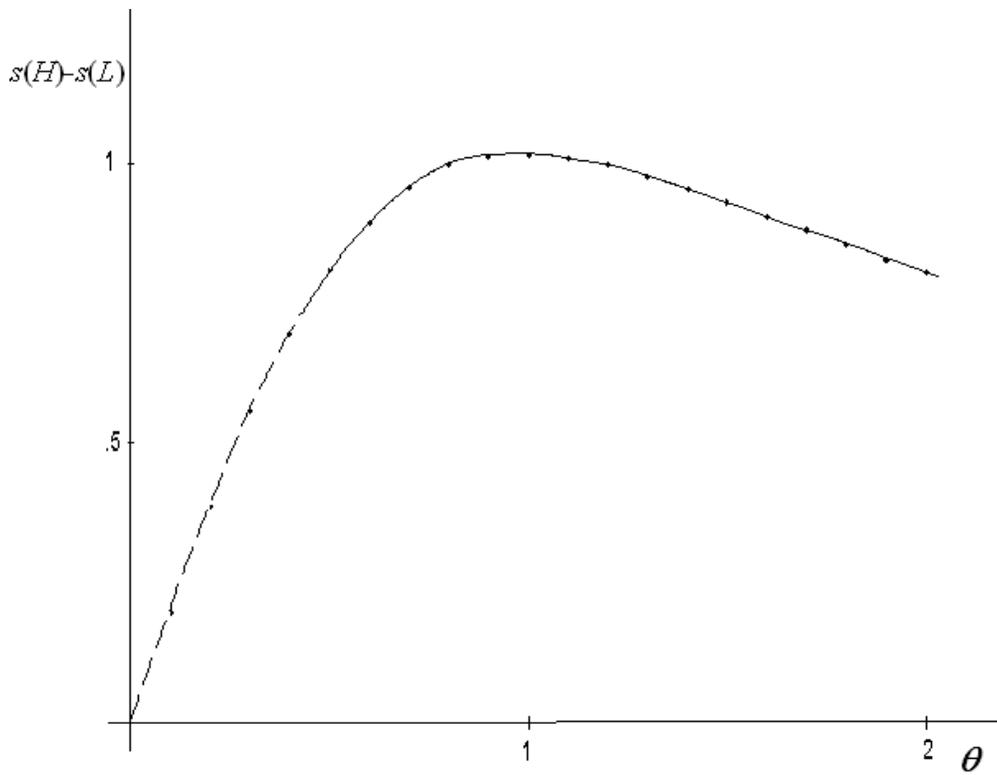


Figure 4: Figure 4: Risk and PPS