Bureau
d'économie
théorique
et appliquée
(BETA)
UMR 7522

Documents de travail

« Incentive Contracts and Downside Risk Sharing »

Auteurs

Bernard Sinclair-Desgagné, Sandrine Spaeter

Document de Travail n° 2016 – 22

Avril 2016

Faculté des sciences économiques et de gestion
Pôle européen de gestion et d'économie (PEGE)
61 avenue de la Forêt Noire
F-67085 Strasbourg Cedex

Secrétariat du BETA
Géraldine Del Fabbro
Tél. : (33) 03 68 85 20 69
Fax : (33) 03 68 85 20 70
g.delfabbro @unistra.fr
www.beta-umr7522.fr

Incentive Contracts and Downside Risk Sharing*

Bernard Sinclair-Desgagné† Sandrine Spaeter‡

January 13, 2016

Abstract

This paper seeks to characterize incentive compensation in a principal-agent moral hazard setting in which the principal is prudent, or downside risk averse, as many situations (such as that of a patient in hospital or a regulator dealing with food safety) suggest she should be. We show that optimal incentive pay should then be 'approximately concave' in performance, the approximation being closer the more downside risk averse the principal is compared to the agent. Limiting the agent's liability would improve the approximation, but taxing the principal would make it coarser. The notion of an approximately concave function we introduce here to describe the pay-performance relationship is relatively recent in mathematics; it is intuitive and translates into concrete empirical implications, notably for the composition of incentive pay. We also clarify which measure of prudence - among the various ones proposed in the literature - is relevant to investigate the tradeoff between downside risk sharing and incentives.

Keywords: Pay-performance relationship; executive compensation; downside risk aversion; approximate concavity

JEL Classification: D82, M12, M52

*We thank Claude d'Aspremont, Pierre Chaigneau, Bruno Deffains, Dominique Demougin, Frédéric Deroïan, Dominique Henriet, Bertrand Koebel, Henri Loubergé, David Martimort, Frédéric Robert-Nicoud, Harris Schlesinger and Marie-Claire Villeval for their insightful remarks and suggestions. Valuable questions and comments from audiences at the 2012 North American Summer Meeting and the 2015 World Congress of the Econometric Society, CORE-Université Catholique de Louvain, the University of Liverpool, LAMETA-University of Montpellier, the 2014 Canadian Economic Theory conference, École polytechnique-Paris, Aix-Marseille School of Economics (AMSE), the University of Geneva, Washington State University, Paris School of Economics (PSE), the University of Lille, GATE-University of Lyon, and 'Risk and Choice: A Conference in Honor of Louis Eeckhoudt' at the University of Toulouse are also gratefully acknowledged. This work benefitted from financial support through project CSES of the Initiative of Excellence (IDEX) at the University of Strasbourg.

†HEC Montréal, CIRANO and CIRPÉE, 3000 Chemin de la Côte-Sainte-Catherine, Montréal (Québec), Canada H3T 2A7. E-mail: bsd@hec.ca

‡BETA, CNRS and University of Strasbourg, 61 avenue de la Forêt Noire, 67085 Strasbourg, France. E-mail: spaeter@unistra.fr

1 Introduction

Under moral hazard, it is well-known that optimal compensation trades off risk sharing and incentives. The involved parties' risk preferences, particularly their respective risk aversion, therefore matter for explaining the shape of incentive pay. In certain circumstances, one also has to look beyond risk aversion. When an agent is subject to contingent monitoring (Fagart and Sinclair-Desgagné 2007) or a principal-agent relationship is exposed to background risk (Ligon and Thistle 2013), for instance, it has been shown that the agent's prudence (or aversion to downside risk) must be considered.

In many other situations, it seems reasonable to expect the principal's prudence to be relevant. Fear of sickness, for example, is likely to feed a patient's aversion to downside risk as she interacts with her doctor/agent (Courbage and Rey 2006). Prudence might also characterize an insurance customer seeking advice from her broker/agent (Hau 2011). A regulator who adopts a precautionary stand - as it is the case in Europe and the United States in public health, food safety and environmental policy (Barrieu and Sinclair-Desgagné 2006, Wiener et al. 2011) - was shown to exhibit prudent risk preferences (Gollier et al. 2000). A firm's executive who contemplates entry in a foreign market might do so with prudence, since international joint ventures can have major downside risk implications (Reuer and Leiblin 2000). In corporate governance, finally, jurisprudence and the law endow board members with fiduciary duties of loyalty and care towards their corporation (Clark 1985; Gutierrez 2003; Lan and Herakleous 2010; Corporate Law Committee 2011). The duty of care confers board directors a key role in preventing and managing crisis situations (Mace 1971; Williamson 2007; Adams et al 2010).[1] When establishing the CEO/agent's compensation, they should accordingly weigh the risks on the downside differently from those on the upside, thereby exerting "(...) that degree of care, skill, and diligence which an ordinary, *prudent* man would exercise in the management of his own affairs" (Clark 1985, p. 73; emphasis added).

This paper thus examines the shape of incentive compensation in a principal-agent setting in which the principal is prudent, or downside risk averse, as this attribute is currently understood in economics. Formally, someone is prudent when her marginal utility function is strictly convex (Menezes et al. 1980; Kimball 1990). A prudent decision maker dislikes mean *and* variance-preserving transformations that skew the distribution of outcomes to the left (Menezes et al. 1980; Crainich and Eeckhoudt 2008).[2] Equivalently, she prefers additional volatility to be associated with good rather than bad outcomes (Eeckhoudt and Schlesinger 2006; Denuit et al. 2010).

[1] Quoting Adams et al. (2010, p. 58): "People often question whether corporate boards matter because their day-to-day impact is difficult to observe. But when things go wrong, they can become the center of attention. Certainly this was true of the Enron, Worldcom, and Parmalat scandals. The directors of Enron and Worldcom, in particular, were held liable for the fraud that occurred: Enron directors had to pay $168 million to investor plaintiffs, of which $13 million was out of pocket (not covered by insurance); and Worldcom directors had to pay $36 million, of which $18 million was out of pocket."

[2] This suggests that prudence is strictly finer than risk aversion. In recent articles, Crainich et al. (2013) point out that 'even risk lovers can be prudent,' and Deck and Schlesinger (2014) report experimental evidence of this.

Our main result is that incentive compensation should then be 'approximately concave' in performance (in the formal sense due to Páles 2003), the approximation being closer the more prudent the principal is relative to the agent. This proposition sheds light on the tradeoff between downside risk sharing and incentives. Roughly speaking, as they *approach* concavity, incentives become more sensitive to outcomes in the range where they are mediocre, thereby transfering the agent greater downside risk; a more prudent principal would likely afford such incentives. Making this intuition precise, however, required us to first invoke, in a principal-agent moral hazard setting, Modica and Scarsini (2005)'s 'coefficient of downside risk aversion' - instead of Kimball (1990)'s better known 'coefficient of prudence' - and to make use of the relatively recent notion (unused so far in economics and finance) of an approximately concave function.

These findings fit a general research program investigating how an agent's incentive compensation is shaped by the principal's preferences. Our characterization of incentive wages also complements Hemmer et al. (2000)'s and Chaigneau (2015)'s respective findings that relate convexity of the agent's remuneration to the agent's prudence under a risk neutral principal, Hau (2011)'s converse statement that a risk averse principal may 'concavify' the reward function of a risk neutral agent, and Gutierrez Arnaiz and Sallas-Fumás (2008)'s justification of convex-concave bonus schemes under specific outcome distributions. Moreover, describing the pay-performance relationship as *approximately* concave has concrete empirical implications. CEO compensation, for instance, involves capped bonuses and similar concave devices, as well as call options that tend to 'convexify' remuneration; in this context, approaching concavity more closely means that the former elements receive greater weight in the overall pay package.

The rest of the paper unfolds as follows. Section 2 presents the benchmark model - a static principal-agent model where the agent is effort and risk averse while the principal is both risk averse and prudent. Adopting the first-order approach is justified using Jung and Kim (2015)'s recent and very general assumptions. Our central result - that the optimal contract should in that case be *approximately* concave, thereby seeking a balance in the downside risk respectively borne by the agent and the principal - is established in Section 3. The section ends with a pair of numerical examples corroborating this statement. Section 4 looks next briefly into empirical matters. Section 5 extends our main result, showing how two frequent contextual elements - limiting the agent's liability or taxing the principal - might affect the downside-risk/incentives tradeoff: limited liability brings compensation closer to concavity in the range where it varies with performance, taxes have the opposite effect. Section 6 contains concluding remarks. All proofs are in the Appendix.

2 The model

Consider an agent - standing for a physician, a financial advisor, a regulated firm, a foreign subsidiary, a CEO, etc. - whose preferences can be represented by a Von Neumann-Morgenstern utility function $u(\cdot)$

defined over monetary payments. We assume this function is three-times differentiable, increasing and strictly concave, formally $u'(\cdot) > 0$ and $u''(\cdot) < 0$, so the agent is risk averse.

This agent can work for a principal - namely a patient, an individual investor, a regulator, a multinational's executive, a corporate board, etc. - whose preferences are represented by the Von Neumann-Morgenstern utility function $v(\cdot)$ defined over net final wealth. We suppose this function is increasing and strictly concave, i.e. $v'(\cdot) > 0$ and $v''(\cdot) < 0$, so the principal is risk averse. Moreover, the marginal utility $v'(\cdot)$ is convex, i.e. $v'''(\cdot) > 0$, which means that the principal is downside risk averse or (equivalently) prudent (Menezes et al. 1980; Kimball 1990).

The principal's profit depends stochastically on the agent's effort level a. The latter cannot be observed, however, while the agent incurs a cost of effort $c(a)$ that is increasing and convex ($c'(a) > 0$ and $c''(a) \geq 0$). The principal only gets a verifiable signal s, drawn from a compact subset $S = [\underline{s}, \overline{s}]$ of $\mathbb{R}$, which is correlated with the agent's effort through a conditional probability distribution $F(s; a)$. This distribution has a density $f(s; a)$ which is differentiable in a and strictly positive on S. The function $\mathcal{L}_F(s; a) = \frac{f_a(s;a)}{f(s;a)}$ then denotes the likelihood ratio of signal s. Based on the realized value of s, the principal receives a benefit $\pi(s)$ expressed in monetary terms, which we suppose increasing and concave or linear in s ($\pi'(s) > 0$ and $\pi''(s) \leq 0$), and she pays the agent a compensation $w(s)$.

The principal's problem is to find a reward schedule $w(s)$ that maximizes her expected utility, under the constraints that the agent will then maximize his own expected utility (the incentive compatibility constraint) and must *ex ante* receive an expected utility that is not inferior to some external one U_0 (the participation constraint). This can be written formally as

$$\max_{w(s),a} \int_{s \in S} v(\pi(s) - w(s)) dF(s; a) \tag{1}$$

subject to

$$a \in \arg\max_{e} \int_{s \in S} u(w(s)) dF(s; e) - c(e)$$

$$\int_{s \in S} u(w(s)) dF(s; a) - c(a) \geq U_0$$

Without losing generality, we shall concentrate on smooth (i.e., twice continuously differentiable) contracts $w(\cdot)$ such that, for $D_1, D_2, ..., D_n$ a finite partition of S, the derivatives $w'(s) < M_i$ on each D_i, with M_i a positive real number. At any performance signal s, the agent's marginal revenue will therefore be bounded. This amounts to saying that the allowed incentive schemes are locally Lipschitz continuous: for any $w(s)$ on each set D_i, $\mid w(x) - w(s) \mid \leq M_i \mid x - s \mid$ for all $x, s \in D_i$. This restriction, which will be useful later in the proofs, seems rather innocuous, since the exogenous number n and ceilings M_i's are arbitrary, and since any continuous function can be approximated as closely as wanted by a sequence of locally Lipschitz maps (Miculescu 2000).

2.1 The first-order approach

For tractability reasons, one usually replaces the incentive compatibility constraint by a relaxed constraint based on the first-order necessary condition for the agent's utility-maximizing effort a. This transforms the principal's initial problem into the following one:

$$\max_{w(s),a} \int_{s\in S} v(\pi(s) - w(s))dF(s;a) \tag{2}$$

subject to

$$\int_{s\in S} u(w(s))dF_a(s;a) - c'(a) \geq 0, \qquad (\gamma)$$
$$\int_{s\in S} u(w(s))dF(s;a) - c(a) \geq U_0, \qquad (\mu)$$

where γ and μ denote the constraints' respective Lagrange multipliers. This so-called 'first-order approach' delivers for sure a solution to problem (1) under the following hypotheses.

Assumption 1: Under the distribution $F(s;a)$, $\pi(s)$ is such that, for $a_2 > a_1$, $\Pr[\pi(s) \leq \bar{\pi} \mid a_1] \geq \Pr[\pi(s) \leq \bar{\pi} \mid a_2]$ for all $\bar{\pi}$, where the inequality holds strictly on a set of positive measure.

Definition 1 *A subset $Z \subseteq \mathbb{R}^n$ is an increasing set if, for any $z^o \in Z$, any $z \geq z^o$ (i.e., $z_i \geq z_i^o$ for every $i = 1, ..., n$) also belongs to Z.*

Assumption 2: For any increasing set $Z_{\pi\mathcal{L}_F}$ of pairs $(\pi(s), \frac{f_a(s;a)}{f(s;a)})$, $\Pr[s \in T(Z_{\pi\mathcal{L}_F}) \mid a]$ is concave in a, where $T(\cdot)$ is the mapping defined as $T(Z) = \{s \in S \mid (\pi(s), \frac{f_a(s;a)}{f(s;a)}) \in Z\}$, $Z \subseteq \mathbb{R}^2$.

Assumption 1 is a *first-order stochastic dominance* (FOSD) condition with respect to π. It basically says that higher effort by the agent makes higher profits more likely. Assumption 2 is a *concave increasing set probability* (CISP) condition with respect to $(\pi, \mathcal{L}_F)$, which then adds that the agent's effort is subject (probabilistically) to decreasing returns.

The next statement corresponds to Jung and Kim (2015)'s Proposition 11.

Lemma 1 *Problem (2) is equivalent to problem (1) if $F(s;a)$ satisfies Assumptions 1 and 2.*

Jung and Kim (2015) demonstrate this lemma. They also show (Proposition 12) that the assumptions made allow more probability distributions (hence are weaker) than any of those used so far in the literature.

2.2 Downside risk aversion

Let us write $R_u = -\frac{u''}{u'}$ and $R_v = -\frac{v''}{v'}$ the Arrow-Pratt measures of absolute risk aversion corresponding to the agent's and the principal's utility functions u and v respectively. Measuring prudence now brings about some subtleties. Denote $P_u = -\frac{u'''}{u''}$, $P_v = -\frac{v'''}{v''}$ the well-known, and somewhat natural, indices of absolute prudence introduced by Kimball (1990). Our results, however, rather rely on the products $P_u R_u = \frac{u'''}{u'} = d_u$ and $P_v R_v = \frac{v'''}{v'} = d_v$.[3]

[3] This was actually seen (but not spelled out) by Belhaj et al. 2014, in a related but different context.

The coefficients d_u and d_v have been proposed by Modica and Scarsini (2005) to capture the intensity of downside risk aversion (see also Jindapon and Neilson 2007). While P_u and P_v indicate the agent's and the principal's respective propensity to prepare in the face of an *inevitable* risk, d_u and d_v rather show how much the agent and the principal would respectively pay to *avoid* a risk with greater negative skewness. In other words, P_u and P_v deal with the strength of an optimal response to risk; they are attached to a *decision*, and their definition involves establishing a 'precautionary' premium ς via the first-order optimality condition $u'(x-\varsigma) = E[u'(x+\widetilde{\varepsilon})]$, where $E(\widetilde{\varepsilon}) = 0$. By contrast, d_u and d_v have to do with the willingness to possibly avert risk at the outset; they capture an *attitude*, and their definition uses instead the usual equality for establishing a risk premium ξ, $u(x-\xi) = E[u(x+\widetilde{\varepsilon})]$, under mean and variance-preserving lottery transformations. As explained by Crainich and Eeckhoudt (2008), the higher d_u, the larger the compensation that the agent requires to be inflicted additional risk at unfavorable outcomes. Trading off downside risk sharing and incentives, while managing agency costs, therefore rests on the relative magnitudes of d_u and d_v.[4]

Let $Dom(u)$ and $Dom(v)$ refer to the respective domains of the utility functions u and v. We now introduce a key definition.

Definition 2 *For some constant real number $k \geq 1$, the principal is said to be more downside risk averse than the agent by a factor k if, for any real numbers $x \in Dom(u)$ and $y - x \in Dom(v)$, we have that $k \cdot d_u(x) \leq d_v(y-x)$.*

In other words, the principal is more downside risk averse than the agent by a factor $k \geq 1$ if, for any amount y to be split between the two and any agent's share x of this amount, the principal's *coefficient of downside risk aversion* d_v taken at her wealth level $(y-x)$ is at least k times bigger than the agent's own coefficient $d_u(x)$. This definition is trivially met when the agent's utility function is quadratic, since $u'''(x) = 0$ and the agent in this case displays no downside risk aversion ($d_u(x) = 0$). The definition also holds when the agent and the principal respectively have constant absolute risk aversion (CARA) utility functions $u(x) = -\exp(-x)$ and $v(y-x) = -\exp(-k\cdot(y-x))$, with $k \geq 1$, because the coefficients of downside risk aversion are then constant. One more example is when the parties respectively have constant relative risk aversion (CRRA) utility functions $u(x) = \frac{x^{1-\alpha}}{1-\alpha}$ and $v(y-x) = \frac{(y-x)^{1-k\alpha}}{1-k\alpha}$, with $0 < \alpha < 1$, $k \geq 1$, $x \in [b_1, b_2]$, $b_1 > 0$ and $y - x \in (0, b_2]$,. The coefficients are then $d_u(x) = \frac{(1+\alpha).\alpha}{x^2}$ and $d_v(y-x) = \frac{(1+k\alpha)\cdot k\alpha}{(y-x)^2}$, so for all x, $y-x$ we have that $k \cdot d_u(x) \leq d_v(y-x)$ with $k = \max[1, \frac{(1+\alpha)b_2^2 - b_1^2}{\alpha\, b_1^2}]$.

This completes the model's description. Let us now proceed to characterize the optimal incentive scheme in this context.

[4] Incidentally, note that neither the prudence coefficient indices P_u, P_v nor the downside risk aversion coefficients d_u, d_v retain similar global properties as the corresponding Arrow-Pratt measures of risk aversion. Keenan and Snow (2002) propose instead an alternative index which does increase under monotonic downside risk averse transformations of the utility function. One can show, nevertheless, that d_u, d_v increase as their respective utility functions u, v become more concave while the marginal utilities u', v' are more convex (Crainich and Eeckhoudt 2008).

3 The optimal contract

This section will establish that a principal who is more downside risk averse than the agent should set an incentive compensation package that is *approximately* concave in outcome. We shall first define approximate concavity, then state and prove our central result, then illustrate it with two examples.

3.1 Approximate concavity

The following definition is adapted from Páles (2003).[5]

Definition 3 *Let I be a subinterval of the real line $\mathbb{R}$ and δ, ρ some nonnegative real numbers. A function $g : I \longrightarrow \mathbb{R}$ is called (δ, ρ)-concave on I if $tg(x) + (1-t)g(y) \leq g(tx + (1-t)y) + \delta t(1-t) \mid x - y \mid + \rho$ for all $x, y \in I$ and $t \in [0, 1]$.*

The function g is of course concave when $\delta = \rho = 0$. Examples of approximately concave functions are pictured in subsection 3.3 below.

Combining Páles (2003)'s theorems 4 and 5 provides a natural description of a $(\delta, 0)$-concave function, which is the situation we will encounter here.

Lemma 2 *Let I be a subinterval of the real line $\mathbb{R}$ and δ a nonnegative number. A function $g : I \longrightarrow \mathbb{R}$ is $(\delta, 0)$-concave on I if there exists a nonincreasing function $q : I \longrightarrow \mathbb{R}$ such that $g(y) \leq g(x) + q(x)(y - x) + \frac{\delta}{2} \mid y - x \mid$ for all $x, y \in I$.*

For completeness, a proof of this lemma is given in the Appendix.[6] One may notice that the above function q bears a close resemblance to a subgradient. The literature indeed says that g is nonincreasingly $(\delta, 0)$-*subdifferentiable* when such a function exists.

We are now ready to move on.

3.2 *Approximately* concave incentive schemes

In order to characterize the agent's incentive contract, we need to make the following assumption, which implies Assumption 1 when the signal s is one-dimensional (Whitt 1980; Sinclair-Desgagné 1994).

Assumption 3: The likelihood ratio $\mathcal{L}_F(s; a) = \frac{f_a(s;a)}{f(s;a)}$ is nondecreasing and concave in s for any a.

This is the so-called *concave monotone likelihood ratio property* (CMLRP), an assumption which is rather common in principal-agent analysis. It is satisfied by many familiar distributions, such as the Poisson

[5] Páles (2003) actually defines and works with approximately *convex* functions. One has to be careful, because some of his results do not straightforwardly carry over after reversing the inequality sign.

[6] It may be worth mentioning the reader that the direct converse of Lemma 2 is not true, so this lemma does not convey an if-and-only-if characterization of $(\delta, 0)$-concave functions. See Remark 3 in Páles (2003, p. 250) for details.

with mean a, the gamma with mean κa, and the chi-squared with degree of freedom parameter a.[7] It "(...) suggests that variations in output at higher levels are relatively less useful in providing 'information' on the agent's effort than they are at lower levels of output" (Jewitt 1988, p. 1181). Let us stress that assuming a concave likelihood ratio does not make the optimal contract trivially concave in outcome. Hemmer et al. (2000), for instance, need Assumption 3 to justify the first-order approach; yet, their analysis supports the use of *convex* incentive devices.

Now, the Kuhn-Tucker necessary and sufficient conditions require that a solution to program (2) meet the equation

$$\frac{v'(\pi(s)-w(s))}{u'(w(s))} = \mu + \gamma \frac{f_a(s;a)}{f(s;a)}, \qquad \forall s \tag{3}$$

The multiplier γ being positive, Assumption 3 entails that the right-hand-side of (3) is increasing in the signal s. This allows to say the following.

Lemma 3 *The optimal reward schedule $w^*(s)$ is increasing in the performance signal s.*

To prove this lemma, in the Appendix, we take the first derivative with respect to s of the left-hand-side of expression (3), then seek conditions which are necessary to make it positive, since the right-hand side's derivative with respect to s is positive by Assumption 3. Similarly taking the second derivative of expression (3)'s left-hand-side and using the same line of argument yields the central result of this paper.

Theorem 1 *Suppose that the principal is more downside risk averse than the agent by a factor k.*

(i) Then the optimal wage schedule $w^*(s)$ is $(\delta(k),0)$-concave at any $s \in S$.

(ii) The number $\delta(k)$ decreases with k and tends to 0 as k grows.

The proof in the Appendix shows that convergence to concavity may not be asymptotic: when $\left(\frac{\pi'(s)-w'(s)}{w'(s)}\right)^2 \geq \frac{1}{k}$, i.e. the agent's earnings do not grow too fast with respect to the principal's net benefit, then we have $\delta(k) = 0$ so $w^*(s)$ is concave at s. In the region where it is not perfectly concave, moreover, the optimal wage schedule has a nonincreasing $(\delta(k),0)$-subgradient which is precisely the derivative $\pi'(s)$ of the principal's benefit function; the contracted payment is thus aligning - albeit imperfectly - the agent's incentives on the principal's interests.

The theorem's conclusion holds vacuously - hence the optimal incentive scheme is concave - when the agent is not prudent (since $u''' \leq 0$ implies $d_u \leq 0$, k can then be as big as wanted). If the agent is prudent (i.e. $u''' > 0$), the theorem says that the pay-performance relationship will only get *closer* to concavity as the principal's downside risk aversion grows.[8] For concreteness, the upcoming subsection provides some numerical examples.

[7] These distributions can be transformed into ones that have compact support and satisfy the previous assumptions by suitably reallocating the probability mass in the tails.

[8] To be very clear, the term 'closer' we have used so far, and will be using later on, does *not* refer to greater curvature, but to being more like *some* concave function (which turns out here to be an affine transformation of the profit function).

3.3 Two examples

Let the principal's benefit function take the specific form $\pi(s) = \sqrt{s}$, $s \in [0, +\infty)$, and let the distribution $F(s; a)$, parameterized by the agent's discrete effort levels $a \in \{1, 2, 3\}$, be a gamma distribution with mean κa, $0 < \kappa$.

■ *The agent is not prudent*

First, consider a risk averse and prudent principal and a risk averse but non-prudent agent. Let their respective utility functions be given by $u(x) = x - bx^2$ and $v(z) = \ln(z+10)$, taking $b > 0$ so that $u'(x) > 0$ on the relevant domain. The first-order condition (3) comes up to

$$\begin{aligned} \frac{(\sqrt{s} - w^*(s) + 10)^{-1}}{1 - 2bw^*(s)} &= A(s) \\ 1 &= A(s).(1 - 2bw^*(s)).(\sqrt{s} - w^*(s) + 10) \\ 0 &= 2bw^{*2}(s) - (1 + 2b.(\sqrt{s} + 10)).w^*(s) + (\sqrt{s} + 10 - A^{-1}(s)) \end{aligned} \tag{4}$$

where $A(s) = \mu + \gamma \frac{s - \kappa a}{a^2}$. Hence, the optimal reward function is[9]

$$w^*(s) = \frac{B(s) - \sqrt{B(s)^2 - 8b(\sqrt{s} + 10 - A(s)^{-1})}}{4b} \tag{5}$$

with $B(s) = (1 + 2b.(\sqrt{s} + 100))$. It can be checked that this increasing reward function satisfies $w^{*\prime\prime}(s) < 0$, so the optimal incentive scheme is concave, as expected. This scheme is shown in Figure 1, using parameters $\mu = 0.2, \gamma = 0.1, \kappa = 1, b = 0.035$. The optimal level of effort is $a^* = 2$; it is obtained by introducing (5) in the objective function of the agent and by solving Program (2). The scheme is indeed quite steep over the region of lower outcomes, so the non-prudent agent bears significant downside risk.

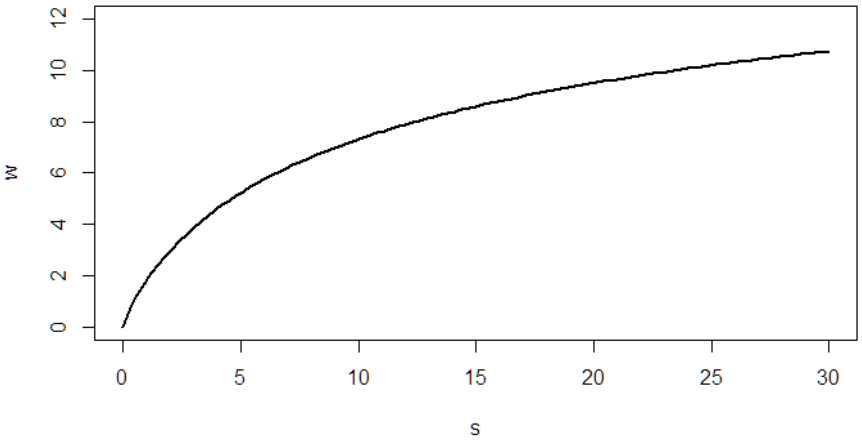

Figure 1. $w^*(s)$ when the agent is risk averse but non prudent ($u'''(\cdot) = 0$)

[9] Two roots exist. We choose the one that satisfies the property $w'(s) > 0$ in optimum. The other one is decreasing in s.

■ *The principal and the agent have CRRA utility functions*

Now, suppose the agent and the principal have respective CRRA utility functions $u(x) = \frac{x^{1-\alpha}}{1-x}$ and $v(z) = \frac{z^{1-k\alpha}}{1-k\alpha}$, with $0 < \alpha < 1$ and $k \geq 1$.

For $k = 1$, the first-order condition (3) becomes

$$\left(\frac{\sqrt{s} - w(s)}{w(s)}\right)^{-\alpha} = \mu + \gamma \frac{s - \kappa a}{a^2}, \qquad \forall s \ .$$

Write again $A(s) = \mu + \gamma \frac{s-\kappa a}{a^2}$. We have that

$$w_1^*(s) = \frac{\sqrt{s}}{1 + A(s)^{-1/\alpha}} \ . \tag{6}$$

When $k = 2$, condition (3) amounts instead to

$$\left(\sqrt{s} - w(s)\right)^2 = A(s)^{-1/\alpha} \cdot w(s), \qquad \forall s \ .$$

The latter equation has one increasing in s root

$$w_2^*(s) = \frac{(2\sqrt{s} + A(s)^{-1/\alpha}) - \sqrt{(2\sqrt{s} + A(s)^{-1/\alpha})^2 - 4s}}{2} \ . \tag{7}$$

Figure 2 portrays the two incentive schemes (6) and (7), using the values $\mu = 0.2$, $\gamma = 0.1$, $\kappa = 1$, and $\alpha = 0.5$. The bold (respectively soft) curve is the optimal reward function obtained in the case $k = 1$ (respectively $k = 2$) and with optimal level of effort $a^* = 2$ (respectively $a^* = 3$). One observes that the two schemes are convex-concave, with the scheme where $k = 2$ approaching a concave function more closely, as the Theorem says.

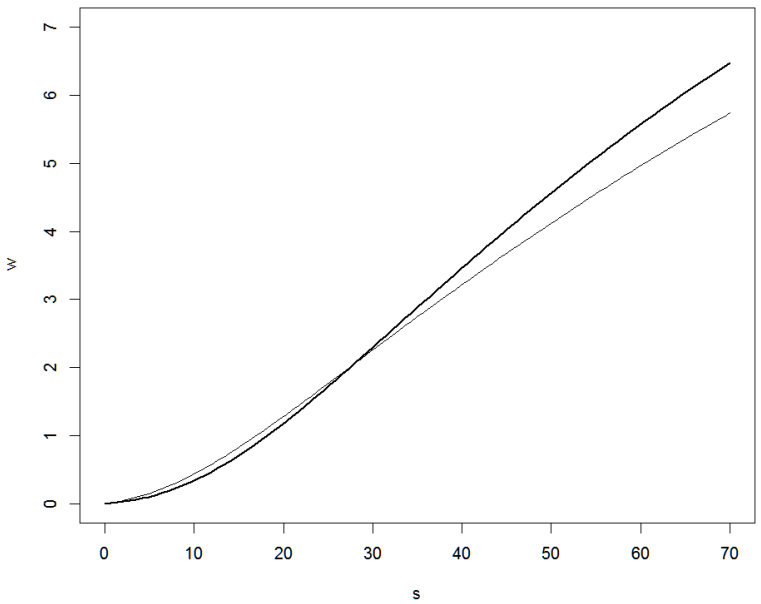

Figure 2. The incentive scheme for CRRA utility functions, when $k = 1$ (bold line) and $k = 2$ (soft line)

4 Empirical considerations

An overall concern with principal-agent theory is whether the predicted incentive contracts can be tested empirically (see, e.g., Prendergast 2002, Chiappori and Salanié 2003, among others). Notwithstanding the standard hurdles (endogeneity problems, notably), another one in the present context might relate to the measurement of the above theorem's and propositions' two main ingredients - downside risk aversion and $(\delta, 0)$-concavity. These issues can be addressed as follows.

Rephrasing Páles (2003)'s corollary 4, one can say that a function $g : I \longrightarrow \mathbb{R}$ is (δ, ρ)-concave on I if it can be represented as the sum $g = h + \ell + b$, where $h : I \longrightarrow \mathbb{R}$ is a concave function, $\ell : I \longrightarrow \mathbb{R}$ is $\delta/2$-Lipschitz continuous, i.e. $\mid \ell(x) - \ell(y) \mid \leq \delta/2 \mid x - y \mid$ for all $x, y \in I$, and $b : I \longrightarrow \mathbb{R}$ is bounded $\parallel b \parallel \leq \rho/2$ on I. This proposition suggests and supports a straightforward manner to capture the $(\delta(k), 0)$-convavity of an incentive contract: simply decompose the pay package into its concave component (e.g., downside penalties, put options, or capped linear bonuses) and its non-concave one (call options, for example). The Lipschitz constant attached to the latter component will correspond to half the coefficient $\delta(k)$, and a convenient proxy for this constant might be the weight the contract is putting on nonconcave incentive devices.

Regarding downside risk aversion, one may want to measure the agent's and the principal's respective indices d_u and d_v directly or, knowing that $d_u = P_u R_u$ and $d_v = P_v R_v$, assess the agent's and the principal's respective degrees of risk aversion R_u, R_v and prudence P_u, P_v by applying some measurement method or using already available data.[10] Note that a principal's prudence might sometimes be inferred indirectly. In the vast empirical literature on CEO compensation, for instance, certain works suggest that regulation (in the utility sector, notably) or public pressure (especially for 'socially responsible' firms) would add to fiduciary duties to ultimately make corporate boards/principals quite prudent.[11]

To test the above statements, moreover, what matters is not only whether the principal is downside risk averse but also, and more importantly, how much more downside risk averse she is relative to the agent. A good proxy for this gap might be some estimate of the agent's influence on the principal; the greater it is, the lower the factor k in the above theorem and propositions. Several measures of a CEO/agent's decision making discretion and power have actually been developed (see Adams et al. 2005, section 2, for a useful account). Similar works exist in health economics concerning the balance of power in patient-doctor relationships (e.g., Goodyear-Smith and Buetow 2001; Gans and Leigh 2011).

[10] For an account of the methods and results concerning the measurement of risk aversion and prudence, see, e.g., Eisenhauer and Ventura (2003), Meyer and Meyer (2006), and Deck and Schlesinger (2014).

[11] Indeed, Murphy (1999) for the utility sector, Frye et al. (2006) on socially responsible firms, and Boschen and Smith (2014) considering the long run report findings that tend to corroborate our results.

5 Extensions

An important issue for the theory of incentives is whether, and to what extent, an incentive contract is influenced by its contextual background. Two customary elements of this background are limited liability and taxation. To check the robustness of our characterization, and thereby sharpen intuition on the tradeoff between downside risk sharing and incentives, this section will now successively consider these elements. It will be shown that they effectively alter the tradeoff, but in opposite ways.

5.1 Limiting the agent's liability

Suppose the agent's revenue is bounded from below, so he cannot bear very high penalties when performance is bad. Remuneration is frequently subject to this type of constraint.[12] An agent with limited wealth, for instance, can file for bankruptcy if he cannot afford paying some penalty. In other contexts, institutions that prevent an agent from breaching his contract under bad circumstances might simply not exist.

Without loss of generality, let us then normalize the agent's minimum revenue to zero. The principal's optimization problem becomes

$$\max_{w(s),a} \int_{s\in S} v(\pi(s)-w(s))dF(s;a) \tag{8}$$

subject to

$$\begin{aligned}
&\int_{s\in S} u(w(s))dF_a(s;a) - c'(a) \geq 0, \qquad (\gamma_1)\\
&\int_{s\in S} u(w(s))dF(s;a) - c(a) \geq U_0, \qquad (\mu_1)\\
&w(s) \geq 0, \forall s \qquad (\lambda(s))
\end{aligned}$$

where $\lambda(s)$ is the Lagrange multiplier associated with the nonnegative wage constraint at signal s.

The Kuhn-Tucker conditions for a solution to this problem are this time given by the equation

$$\frac{v'(\pi(s)-w(s))}{u'(w(s))} = \mu_1 + \gamma_1 \frac{f_a(s;a)}{f(s;a)} + \frac{\lambda(s)}{f(s;a)u'(w(s))}, \qquad \forall s \tag{9}$$

with $\lambda(s)w(s) = 0$ at all s. Arguments and computations similar to those of Section 3 lead to the following statement.

Proposition 1 *Suppose that the principal is more downside risk averse than the agent by a factor k. If the agent is protected by limited liability, then:*

[12] Hence, since Holmstrom (1979) and especially Sappington (1983)'s seminal works, analyzing the impact of an agent's limited liability remains a rather well-covered topic in the principal-agent literature. For a recent account of this literature, see Poblete and Spulber (2012). In most articles, both the principal and the agent are assumed to be risk neutral.

(i) The optimal wage schedule is such that $w_1(s) = 0$ *for any signal* s *below a threshold* s_1 *and* $w_1(s)$ *is* $(\delta_1(k), 0)$*-concave when* $s > s_1$*.*

(ii) The number $\delta_1(k)$ *decreases with* k *and tends to* 0 *as* k *grows.*

(iii) For $s > s_1$ *and any given* k*,* $\delta_1(k) \leq \delta(k)$ *so* $w_1(s)$ *is closer to being concave than the incentive scheme* $w^*(s)$ *obtained in Theorem 1.*

On the range where pay varies with performance, the proposition describes an optimal incentive scheme analogous to the one outlined in the above Theorem. But part (iii) adds an intuitive feature. Since limited liability shelters the agent from the worst outcomes, the prudent principal can afford having him bear more downside risk across the range where compensation is linked to outcomes. On this range, she then selects a wage schedule that is *closer* to a concave one than the schedule she would choose under no limited liability constraint.

5.2 Taxing the principal's benefits

Suppose now that the principal's net benefits are subject to a constant tax rate θ which applies when they are positive. The incentive scheme chosen by the principal must be a solution to

$$\max_{w(s),a} \int_{s\in\overline{S_2}} v((1-\theta)(\pi(s)-w(s)))dF(s;a) + \int_{s\in\underline{S_2}} v(\pi(s)-w(s))dF(s;a) \tag{10}$$

subject to

$$\int_{s\in S} u(w(s))dF_a(s;a) - c'(a) \geq 0, \qquad (\gamma_2)$$
$$\int_{s\in S} u(w(s))dF(s;a) - c(a) \geq U_0, \qquad (\mu_2)$$

where $\overline{S_2} = \{s \in S; \pi(s) - w(s) \geq 0\}$ and $\underline{S_2} = \{s \in S; \pi(s) - w(s) < 0\}$ are endogenous sets.

The Kuhn-Tucker conditions in this case are given by two equations:

$$\frac{(1-\theta)v'\left((1-\theta)(\pi(s)-w(s))\right)}{u'(w(s))} = \mu_2 + \gamma_2 \frac{f_a(s;a)}{f(s;a)}, \quad \forall s \in \overline{S_2} \ , \tag{11}$$

$$\frac{v'(\pi(s)-w(s))}{u'(w(s))} = \mu_2 + \gamma_2 \frac{f_a(s;a)}{f(s;a)}, \quad \forall s \in \underline{S_2} \ . \tag{12}$$

The next proposition, which is derived using the same proof arguments as before, expresses how the principal then trades off downside risk sharing and incentives.

Proposition 2 *Suppose that the principal is more downside risk averse than the agent by a factor* k*. If a constant tax rate* θ *applies to the principal's net benefits when they are positive, then:*

(i) the optimal wage schedule is $(\delta_2(k), 0)$*-concave when net benefits are negative, where the number* $\delta_2(k)$ *gets smaller with* k *and tends to* 0 *as* k *grows;*

(ii) the optimal wage schedule is $(\delta_2(k,\theta),0)$*-concave when net benefits are positive, where the number* $\delta_2(k,\theta)$ *gets smaller with* k *and tends to* 0 *as* k *grows;*

(iii) $\delta_2(k,\theta)$ *increases with* θ*, so the higher the tax rate the cruder the pay-performance concavity.*

Statements (i) and (ii) characterize the chosen incentive scheme over the complementary domains where net benefits are respectively negative and positive; in both cases, the optimal wage schedule mirrors the one described in Theorem 1. Part (iii), however, raises the possibility of incentive pay leaning away from concavity. Recall that downside risk aversion makes the principal worry more about the variability of net benefits in bad states than in good states. When positive net benefits are taxed, the principal associates positive net benefits with states which are relatively less good; she then becomes more sensitive to risk at those states. As the tax rate increases, willingness to have the agent bear more of this risk grows, so the pay-performance relationship turns further aside from concavity.

6 Concluding remarks

When the principal or the agent are prudent (or downside risk averse), an optimal contract must trade off *downside* risk sharing and incentives. This paper first pointed out that the appropriate measure of downside risk aversion to investigate this tradeoff is the coefficient introduced by Modica and Scarsini (2005). This coefficient - the ratio of the third over the *first* derivative of the utility function - precisely indicates how much someone would be willing to pay to avoid a risk with greater negative skewness. Next, we showed that an optimal incentive package will be *approximately* concave in the relevant outcomes (in the precise mathematical sense due to Páles 2003), the approximation being closer the more the principal is downside risk averse compared with the agent. Intuitively, *approaching* a concave function shifts more downside risk upon the agent, since remuneration is then more variable *on the whole* in adverse circumstances. This result is qualitatively robust to limiting the agent's liability or taxing the principal, although these common contextual features might affect the approximation in opposite directions (making it better or coarser, respectively).

Our characterization of incentive pay adds to the literature in one more aspect. Previous investigations of the pay-performance relationship have seeked to support or explain the use of a peculiar device, like stock options or discretionary bonuses. In practice, however, incentive contracts are compound packages made of various items, some concave and other convex or linear in outcome. As we argued in Section 4, the above results - thanks to the notion of *approximate* concavity - can cope with such packages: coming closer to concavity means that convex devices like call options have become relatively less significant in the overall pay package than concave items like linear bonuses with a cap. Our theorem and propositions thus have strong empirical content; meeting the data seems doable and is certainly looked out for. This addresses a recurrent concern with principal-agent theory.

All in all, finally, this paper can be seen as fitting a research program that studies how compensation is shaped by plausible attributes of the principal's preferences. Pursuing such a program further might deliver other empirically testable insights on remuneration. It would also allow, in some cases, a *normative* analysis of incentive pay. In some situations, indeed, certain risk attitudes by the principal may be called for. In health care, for instance, one may take the Hyppocratic oath as a traditional call for medical prudence (see Linden 1999). In corporate governance, one may interpret the fiduciary duty of care that corporate boards/principals must obey as imposing 'prudence' on their decision making (see Clark 1985), in particular when they set executive pay. In these contexts, statements like the ones contained in our theorem and propositions would have to be seen as *requirements* (not predictions) on the pay-performance relationship.

APPENDIX

Proof of Lemma 2.

Since we work here with *approximately concave* rather than approximately convex functions, the proof combines and adapts our notation and context to the arguments underlying theorems 4 and 5 in Páles 2003).

Suppose there is a nonincreasing function $q : I \longrightarrow \mathbb{R}$ such that $g(y) \leq g(x) + q(x)(y-x) + \frac{\delta}{2} \mid y-x \mid$ for all $x, y \in I$. The latter inequality can be rewritten as

$$q(x)(y-x) \geq g(y) - g(x) - \frac{\delta}{2} \mid y-x \mid \quad \text{for all} \quad x, y \in I \ .$$

Take the primitive function of q, which is $Q(x) = \int_{x_0}^{x} q(t)dt$ where x_0 is an arbitrary fixed element of I. This function $Q : I \to \mathbb{R}$ is concave since q is non increasing.

Let $x = t_0 < t_1 < ... < t_n = y$ be an arbitrary grid on I. Applying the above inequality successively to t_{i-1} and t_i for $i = 1, ..., n$ and summing up gives

$$\begin{aligned} \sum_{i=1}^{n} q(t_i)(t_{i-1} - t_i) &\geq \sum_{i=1}^{n} \left(g(t_{i-1}) - g(t_i) - \frac{\delta}{2}(t_i - t_{i-1}) \right) \\ &= g(x) - g(y) - \frac{\delta}{2}(y-x) \ . \end{aligned}$$

If $\max_i \ (t_i - t_{i-1})$ is small enough, the latter implies that

$$Q(x) - Q(y) \geq g(x) - g(y) - \frac{\delta}{2}(y-x) \ .$$

Similarly, apply the above inequality to t_i and t_{i-1} for $i = 1, ..., n$ and take the sum. This yields

$$\begin{aligned}\sum_{i=1}^{n} q(t_{i-1})(t_i - t_{i-1}) &\geq \sum_{i=1}^{n} \left(g(t_i) - g(t_{i-1}) - \frac{\delta}{2}(t_i - t_{i-1}) \right) \\ &= g(y) - g(x) - \frac{\delta}{2}(y - x) \ .\end{aligned}$$

Hence, if the grid is again fine enough, we have that

$$Q(y) - Q(x) \geq g(y) - g(x) - \frac{\delta}{2}(y - x) \ .$$

Define $\ell : I \to \mathbb{R}$ as $\ell(x) = g(x) - Q(x)$. The upshot is that

$$\mid \ell(x) - \ell(y) \mid \ \leq \frac{\delta}{2}(y - x) \ \text{ for all } \ x < y \ \text{ in } \ I \ .$$

For $t \in [0, 1]$, this conclusion and the triangle inequality entail that

$$\begin{aligned}t\ell(x) + (1-t)\ell(y) &= \ell(tx + (1-t)y) + t(\ell(x) - \ell(tx + (1-t)y)) + (1-t)(\ell(y) - \ell(tx + (1-t)y)) \\ &\leq \ell(tx + (1-t)y) + \mid t(\ell(x) - \ell(tx + (1-t)y)) + (1-t)(\ell(y) - \ell(tx + (1-t)y)) \mid \\ &\leq \ell(tx + (1-t)y) + t \mid \ell(x) - \ell(tx + (1-t)y) \mid + \ (1-t) \mid \ell(y) - \ell(tx + (1-t)y) \mid \\ &\leq \ell(tx + (1-t)y) + t\frac{\delta}{2}(1-t) \mid x - y \mid + (1-t)\frac{\delta}{2}t \mid x - y \mid \\ &= \ell(tx + (1-t)y) + t(1-t)\delta \mid x - y \mid \ \text{ for all } \ x, y \in I \ .\end{aligned}$$

Now, since Q is concave, we have $tQ(x) + (1-t)Q(y) \leq Q(tx + (1-t)y)$. Adding this inequality and the latter one leads to

$$tg(x) + (1-t)g(y) \leq g(tx + (1-t)y) + \delta t(1-t) \mid x - y \mid \ \text{ for all } \ x, y \in I \ \text{ and } \ t \in [0, 1] \ .$$

Hence, g is $(\delta, 0)$-concave. ♦

Proof of Lemma 3.

Risk aversion of at least one player is sufficient to obtain that $w^{*\prime}(s) \geq 0$. Indeed we have, with $v(\pi(s) - w(s))$ denoted as $v(\cdot)$ and $u(w(s))$ denoted as $u(\cdot)$:

$$\begin{aligned}&\frac{\partial}{\partial s}\left(\frac{v'(\pi(s) - w(s))}{u'(w(s))} \right) \\ = \ &\frac{(\pi'(s) - w'(s)).v''(\cdot)u'(\cdot) - v'(\cdot).u''(\cdot).w'(s)}{(u'(\cdot))^2} \\ = \ &\frac{-w'(s).\,(v''(\cdot).u'(\cdot) + v'(\cdot).u''(\cdot)) + \pi'(s).v''(\cdot).u'(\cdot)}{u'(\cdot)^2}\end{aligned}$$

The latter must be positive, by Assumption 3, in order to satisfy equation (3). A necessary condition for this is $w'(s) \geq 0$. ♦

Proof of Theorem 1.

Let us now compute the second derivative of the left-hand side term in equation (3). Assumption 3 entails it must be negative.

$$\begin{aligned}
&\frac{\partial^2}{\partial s^2}\left(\frac{v'(\pi(s)-w(s))}{u'(w(s))}\right)\\
=\;& \frac{1}{(u'(.))^4}.\left(\{-w''.(v''u'+v'u'')-w'.\left[(\pi'-w')v'''u'+w'v''u''+(\pi'-w')v''u''+w'v'u'''\right]\right.\\
&\left.+\pi''v''u'+\pi'\left[(\pi'-w')v'''u'+v''w'u''\right]\}.u'^2+2u''u'.w'.\left[w'(v''u'+v'u'')-\pi'v''u'\right]\right)\\
=\;& \frac{1}{u'^3}.\left(\left\{-w''.(v''u'+v'u'')+(\pi'-w')^2v'''u'+(\pi'-w')w'v''u''-w'^2v'u'''+\pi''v''u'\right\}.u'\right.\\
&\left.-(\pi'-w')v''u''w'u'(u'+2)+2u''^2w'^2v'\right)\\
=\;& \frac{1}{u'^3}.\left(\left\{-w''.(v''u'+v'u'')+(\pi'-w')^2v'''u'-w'^2v'u'''+\pi''v''u'\right\}.u'\right.\\
&\left.-2w'u''.\left((\pi'-w')v''u'-v'u''w'\right)\right.\\
=\;& \frac{1}{u'^2}.\left[-w''.(v''u'+v'u'')+(\pi'-w')^2v'''u'-w'^2v'u'''+\pi''v''u'\right]\\
&+2w'R_u.\frac{\left((\pi'-w')v''u'-v'u''w'\right)}{u'^2}
\end{aligned}$$

The last term here is in fact $\frac{\partial}{\partial s}\left(\frac{v'(\pi(s)-w(s))}{u'(w(s))}\right)$, which must be positive by equation (3) and Assumption 3. Then:

$$\begin{aligned}
&\frac{\partial^2}{\partial s^2}\left(\frac{v'(\pi(s)-w(s))}{u'(w(s))}\right)\\
=\;& 2w'R_u.\frac{\partial}{\partial s}\left(\frac{v'(\pi(s)-w(s))}{u'(w(s))}\right)\\
&+\frac{1}{u'^2}.\left[-w''.(v''u'+v'u'')+(\pi'-w')^2v'''u'-w'^2v'u'''+\pi''v''u'\right]\\
=\;& 2w'R_u.\frac{\partial}{\partial s}\left(\frac{v'(\pi(s)-w(s))}{u'(w(s))}\right)\\
&+\frac{v'}{u'}.\left[w''.(R_v+R_u)+(\pi'-w')^2P_vR_v-w'^2P_uR_u-\pi''R_v\right]\\
=\;& 2w'R_u.\frac{\partial}{\partial s}\left(\frac{v'(\pi(s)-w(s))}{u'(w(s))}\right)\\
&+\frac{v'}{u'}.\left[w''.(R_v+R_u)-\pi''R_v+(\pi'-w')^2P_vR_v-w'^2P_uR_u\right]
\end{aligned}$$

The sign of this last expression depends on the sign of $(\pi'-w')^2P_vR_v-w'^2P_uR_u$, which writes explicitly as

$$\begin{aligned}
&(\pi'(s)-w'(s))^2P_v(\pi(s)-w(s))R_v(\pi(s)-w(s))-w'(s)^2P_u(w(s))R_u(w(s))\\
=\;& (\pi'(s)-w'(s))^2\cdot d_v(\pi(s)-w(s))-w'(s)^2\cdot d_u(w(s))\ .
\end{aligned}$$

Two cases are possible. Either

$$\left(\frac{\pi'(s)-w'(s)}{w'(s)}\right)^2\geq\frac{1}{k}\ ,$$

and

$$\frac{1}{k} \geq \frac{d_u(w(s))}{d_v(\pi(s) - w(s))}$$

by assumption and Definition 1 makes it necessary that $w''(s) < 0$ (so w is concave at s); or

$$\left(\frac{\pi'(s) - w'(s)}{w'(s)}\right)^2 < \frac{1}{k} \quad . \tag{13}$$

Recall that the derivatives $w'(s)$ are locally bounded by assumption. Thus, there exist positive real numbers M_i associated with each D_i, $i = 1..n$, of a finite partition $D_1, D_2, ..., D_n$ of S such that $w'(s) < M_i$ on D_i. Let $M = \max_i M_i$ and take $\delta(k) > 0$ so that $\left(\frac{\delta}{4M}\right)^2 = \frac{1}{k}$. Inequality (13) is equivalent to

$$(\pi'(s) - w'(s))^2 < \left(\frac{\delta(k)}{4M}\right)^2 (w'(s))^2 \ .$$

For all $x \in S$, $x \neq s$, then:

$$\begin{aligned} \mid (\pi'(s) - w'(s))(x - s) \mid \ &< \left(\frac{\delta(k)}{4M}\right) w'(s) \mid x - s \mid \\ &< \quad \frac{\delta(k)}{4} \mid x - s \mid . \end{aligned}$$

Hence,

$$w'(s)(x - s) \ < \pi'(s)(x - s) + \frac{\delta(k)}{4} \mid x - s \mid .$$

Now, since w is differentiable at s, we have that

$$w(x) = w(s) + w'(s)(x - s) + r(x)$$

with the residual $r(x)$ satisfying $\lim_{x \to s} \frac{r(x)}{x - s} = 0$. The last inequality entails that

$$\begin{aligned} w(x) \quad &< \quad w(s) + \pi'(s)(x - s) + \left(\frac{r(x)}{\mid x - s \mid} + \frac{\delta(k)}{4}\right) \mid x - s \mid \\ &\leq \quad w(s) + \pi'(s)(x - s) + \frac{\delta(k)}{2} \mid x - s \mid \end{aligned}$$

if x is sufficiently close to s. Since $\pi'(s)$ is decreasing in s by assumption, applying Lemma 2 yields that $w(s)$ is $(\delta(k), 0)$-concave on a subinterval of S that contains s. Since this is to be true at any point s, keeping the same number $\delta(k)$, then $w^*(s)$ is $(\delta(k), 0)$-concave on S. This proves assertion (i).

Assertion (ii) is immediate when considering the above construction of $\delta(k)$, as $\left(\frac{\delta}{4M}\right)^2 = \frac{1}{k}$. ♦

Proof of Proposition 1.

The conclusion of Lemma 3 still holds here, so the optimal wage schedule $w_1(s)$ must be increasing in its argument s. There is therefore a threshold s_1 such that $w_1(s) = 0$ for $s \leq s_1$ and $w_1(s) > 0$ for $s > s_1$.

For any signal $s \leq s_1$, the limited liability constraint is binding so $w(s) = 0$. For $s > s_1$, we have $\lambda(s) = 0$; equation (9) is then similar to equation (3), and the proof of the Theorem can be replicated to conclude that $w(s)$ is $(\delta_1(k), 0)$-concave. This proves Assertion (i) of the proposition.

Assertion (ii) is immediate since δ_1 satisfies $\left(\frac{\delta_1}{4M'}\right)^2 = \frac{1}{k}$.

To show (iii), notice that the proof of (i) uses the bound $M' = \max_{D_i \cap S_1 \neq \varnothing} M_i$, considering only the neighborhoods D_i having a nonempty intersection with $S_1 = \{s \in S \ ; \ s > s_1\}$. Clearly, $M' \leq M = \max_{all\ D_i} M_i$. Hence, $\left(\frac{\delta}{4M}\right)^2 = \frac{1}{k} = \left(\frac{\delta_1}{4M'}\right)^2$ implies that $\delta_1(k) \leq \delta(k)$ for any given k. ♦

Proof of Proposition 2.

When $s \in \underline{S_2}$, net benefits are negative so the principal is not taxed. Theorem 1 then applies and the optimal wage schedule $w_2(s)$ is $(\delta_2(k), 0)$-concave at s, where the number $\delta_2(k)$ decreases with k and tends to 0 as k grows.This proves assertion (i).

Now, for any $s \in \overline{S_2}$,

$$\begin{aligned} & \frac{\partial}{\partial s}\left(\frac{v'((1-\theta)(\pi(s)-w(s)))}{u'(w(s))}\right) \\ = & \frac{(1-\theta)(\pi'-w')v''u' - v'u''w'}{u'^2} \\ = & \frac{-w'(1-\theta)(v''u'+v'u'') + (1-\theta)\pi'v''u'}{u'^2} \\ = & \frac{v'(R_u w' - CR_v)}{u'} \end{aligned} \tag{14}$$

where $C(s) = \frac{\partial}{\partial s}\left[(1-\theta)(\pi(s)-w(s))\right] = (1-\theta)(\pi'(s)-w'(s))$. From (14), $w'(s) > 0$ is a necessary condition for $\frac{\partial}{\partial s}\left(\frac{v'((1-\theta)(\pi(s)-w(s)))}{u'(w(s))}\right) > 0$.

Computation of the second derivative gives

$$\begin{aligned} & \frac{\partial^2}{\partial s^2}\left(\frac{v'((1-\theta)(\pi(s)-w(s)))}{u'(w(s))}\right) \\ = & \frac{C'v'' + C^2v''' + \left(R'_u v' + R_u C v''\right)w' + R_u v' w''}{u'} \\ & - \frac{\left(Cv'' + R_u v' w'\right)u''w'}{(u')^2} \\ = & \frac{C'v'' + C^2v''' + \left(R'_u v' + R_u C v''\right)w' + R_u v' w'' + \left(Cv'' + R_u v' w'\right)R_u w'}{u'} \\ = & \frac{v'(C^2 R_v P_v - C'R_v) + \left(R'_u - R_u C R_v\right)v'w' + R_u v' w'' + \left(R_u w' - CR_v\right)R_u w' v'}{u'} \\ = & \frac{v'}{u'}.\left(C^2 R_v P_v - C'R_v + \left(R'_u - R_u C R_v\right)w' + R_u w'' + \left(R_u w' - CR_v\right)R_u w'\right) \\ = & \frac{v'}{u'}.\left[R_v(C^2 P_v - C') + \left(R'_u - R_u C R_v\right)w' + R_u(w'' + \left(R_u w' - CR_v\right)w')\right] \end{aligned} \tag{15}$$

where $\frac{dR_u}{ds} = R'_u = -w'(d_u - R_u^2)$. The component within brackets of Equation (15) reduces to

$$
\begin{aligned}
& R_v(C^2P_v - C') + (R'_u - R_uCR_v)\,w' + R_u(w'' + (R_uw' - CR_v)\,w') \\
= \;& C^2d_v - C'R_v - w'^2d_u + w'^2R_u^2 - R_uCR_vw' + R_uw'.\,(R_uw' - CR_v) + R_u.w'' \\
= \;& \left((1-\theta)^2(\pi' - w')^2.d_v - w'^2d_u\right) + 2w'.R_u.\,(R_uw' - CR_v) - (1-\theta)\pi''R_v + \\
& \qquad ((1-\theta)R_v + R_u).w'' \qquad (16)
\end{aligned}
$$

The second term in (16) is equal to $\frac{2R_uw'(s)u'(s)}{v'(s)}.\frac{\partial}{\partial s}\left(\frac{v'((1-\theta)(\pi(s)-w(s)))}{u'(w(s))}\right)$ which is positive. The third term is positive by assumption. Hence, the sign of $w''(s)$ for any $s \in \overline{S_2}$ depends on the sign of the first term, $(1-\theta)^2(\pi' - w')^2.d_v - w'^2d_u$.

As in the proof of the theorem, two cases are possible. Either

$$\left(\frac{(1-\theta).\,(\pi'(s) - w'(s))}{w'(s)}\right)^2 \geq \frac{1}{k} \quad ,$$

and, since $\frac{1}{k} > \frac{d_u}{d_v}$, the expression $(1-\theta)^2(\pi' - w')^2.d_v - w'^2d_u$ is positive. It is then necessary that $w''(s) < 0$, so w_2 is concave at $s \in \overline{S_2}$. In the opposite situation,

$$\left(\frac{(1-\theta).\,(\pi'(s) - w'(s))}{w'(s)}\right)^2 < \frac{1}{k} \quad .$$

This inequality can be written

$$\left(\frac{\pi'(s) - w'(s)}{w'(s)}\right)^2 < \frac{1}{k\cdot(1-\theta)^2} \quad .$$

Let $M'' = \max\limits_{D_i \,\cap\, \overline{S_2} \neq \varnothing} M_i$, and take $\delta_2(k,\theta) > 0$ so that

$$\left(\frac{\delta_2(k,\theta)}{4M''}\right)^2 = \frac{1}{k\cdot(1-\theta)^2} \quad . \qquad (17)$$

The latter inequality reduces to

$$(\pi'(s) - w'(s))^2 < \left(\frac{\delta_2(k,\theta)}{4M''}\right)^2 (w'(s))^2 \; .$$

For all $x \in \overline{S_2}$, $x \neq s$, it follows that

$$
\begin{aligned}
\mid (\pi'(s) - w'(s))(x-s) \mid \; &< \left(\frac{\delta_2(k,\theta)}{4M''}\right) w'(s) \mid x - s \mid \\
&< \frac{\delta_2(k,\theta)}{4} \mid x - s \mid ,
\end{aligned}
$$

so

$$w'(s))(x-s) \; < \pi'(s)(x-s) + \frac{\delta_2(k,\theta)}{4} \mid x - s \mid .$$

Since w is differentiable at s,

$$w(x) = w(s) + w'(s)(x-s) + r(x)$$

with the residual $r(x)$ such that $\lim_{x \to s} \frac{r(x)}{x-s} = 0$. The last inequality entails that

$$\begin{aligned} w(x) &\leq w(s) + \pi'(s)(x-s) + \left(\frac{r(x)}{\mid x - s \mid} + \frac{\delta_2(k,\theta)}{4} \right) \mid x - s \mid \\ &\leq w(s) + \pi'(s)(x-s) + \frac{\delta_2(k,\theta)}{2} \mid x - s \mid \end{aligned}$$

if x is sufficiently close to s. Since $\pi'(s)$ is decreasing in s by assumption, applying Lemma 2 yields that $w_2(s)$ is $(\delta_2(k,\theta), 0)$-concave in $s \in \overline{S_2}$. This shows assertion (ii).

To demonstrate assertion (iii), note that, as it is defined in (17), $\delta_2(k,\theta)$ must increase with θ. ♦

References

[1] R.B. Adams, H. Almeida, and D. Ferreira. Powerful ceos and their impact on corporate performance. *The Review of Financial Studies*, 18(4):1403–143, 2005.

[2] R.B. Adams, B.E. Hermalin, and M.S. Weisbach. The role of boards of directors in corporate governance: A conceptual framework and survey. *Journal of Economic Literature*, 48:58–107, 2010.

[3] P. Barrieu and B. Sinclair-Desgagné. On precautionary policies. *Management Science*, 52(8):1145–1154, 2006.

[4] M. Belhaj, R. Bourlès, and F. Deroian. Risk-taking and risk-sharing incentives under moral hazard. *American Economic Journal: Microeconomics*, 6(1):58–90, 2014.

[5] J.F. Boschen and K.J. Smith. Asymetries in dynamic ceo compensation. Working Paper, The College of William and Mary, available at SSRN: http://ssrn.com/abstract=2333233 or http://dx.doi.org/10.2139/ssrn.2333233, 2014.

[6] P. Chaigneau. Risk aversion, prudence and compensation. *The European Journal of Finance*, 21:1357–1373, 2015.

[7] P.A. Chiappori and B. Salanié. *Testing Contract Theory: A Survey of Some Recent Work*. In: Advances in Economics and Econometrics, vol 1, M. Dewatripont, L. Hansen and S. Turnovsky eds, Cambridge University Press, 2003.

[8] R.C. Clark. *Agency Costs versus Fiduciary Duties*. In: J.W. Pratt and R.J. Zeckhauser (Ed.), Principals and Agents: the Structure of Business. Harvard Business School Press, 1985.

[9] Corporate Law Committee. *Corporate Director's Guidebook*. American Bar Association Publishing, Chicago IL, 2011.

[10] C. Courbage and B. Rey. Prudence and optimal prevention for health risks. *Health Economics*, 15:1323–1327, 2006.

[11] D. Crainich and L. Eeckhoudt. On the intensity of downside risk aversion. *Journal of Risk and Uncertainty*, 36:267–276, 2008.

[12] D. Crainich, L. Eeckhoudt, and A. Trannoy. Even (mixed) risk lovers are prudent. *American Economic Review*, 103(4):1529–1535, 2013.

[13] C. Deck and H. Schlesinger. Consistency of higher order risk preferences. *Econometrica*, 82(5):1913–1943, 2014.

[14] M. Denuit, L. Eeckhoudt, and B. Rey. Some consequences of correlation aversion in decision sciences. *Annals of Operations Research*, 176:259–269, 2010.

[15] L. Eeckhoudt and H. Schlesinger. Putting risk in its proper place. *American Economic Review*, 96:280–289, 2006.

[16] J.G. Eisenhauer and L. Ventura. Survey measures of risk aversion and prudence. *Applied Economics*, 35:1477–1484, 2003.

[17] M-C. Fagart and B. Sinclair-Desgagné. Ranking contingent monitoring systems. *Management Science*, 53:1501–1509, 2007.

[18] M.B. Frye, E. Nelling, and E. Webb. Executive compensation in socially responsible firms. *Corporate Governance*, 14:446–455, 2006.

[19] J.S. Gans and A. Leigh. Bargaining over labor: Do patients have any power? Discussion Paper series, Forschunginstitut zur Zukunft der Arbeit, No. 6165, 2011.

[20] C. Gollier, B. Jullien, and N. Treich. Scientific progress and irreversibility: An economic interpretation of the 'precautionary principle'. *Journal of Public Economics*, 75:229–253, 2000.

[21] F. Goodyear-Smith and S. Buetow. Power issues in the doctor-patient relationship. *Health Care Analysis*, 9:449–462, 2001.

[22] M. Gutiérrez. An economic analysis of corporate directors' fiduciary duties. *Rand Journal of Economics*, 34:516–535, 2003.

[23] O. Gutiérriez Arnaiz and V. Salas-Fumás. Performance standards and optimal incentives. *Journal of Accounting and Economics*, 45:139–152, 2008.

[24] A. Hau. Optimal brokerage commission for fair insurance: A first order approach. *The Geneva Risk and Insurance Review*, 36:189–201, 2011.

[25] T. Hemmer, O. Kim, and R.E. Verrecchia. Introducing convexity into optimal compensation contracts. *Journal of Accounting and Economics*, 28:307–327, 2000.

[26] B. Holmström. Moral hazard and observability. *The Bell Journal of Economics*, 10:74–91, 1979.

[27] I. Jewitt. Justifying the first-order approach to principal-agent problems. *Econometrica*, 56:1177–1190, 1988.

[28] P. Jindapon and W.S. Neilson. Higher-order generalizations of arrow-pratt and ross risk aversion: A comparative statics approach. *Journal of Economic Theory*, 136:719–728, 2007.

[29] J.Y. Jung and S.K. Kim. Information space conditions for the first-order approach in agency problems. *Journal of Economic Theory*, 160:243–279, 2015.

[30] D. Keenan and A. Snow. Greater downside risk aversion. *The Journal of Risk and Uncertainty*, 24:267–277, 2002.

[31] M. Kimball. Precautionary saving in the small and in the large. *Econometrica*, 58:53–73, 1990.

[32] L.L. Lan and L. Heracleous. Rethinking agency theory: The view from law. *Academy of Management Review*, 35:294–314, 2010.

[33] J.A. Ligon and P. D. Thistle. Background risk in the principal-agent model. *Geneva Risk and Insurance Review*, forthcoming, 2013.

[34] D.E.J. Linden. Gabriele zerbi's de cautelis medicorum and the tradition of medical prudence. *Bulletin of the History of Medicine*, 73(1):19–37, 1999.

[35] M.L. Mace. *Directors : Myth and Reality*. Boston : Harvard Business School Press, 1971.

[36] C. Menezes, C. Geiss, and J. Tressler. Increasing downside risk. *American Economic Review*, 70:921–932, 1980.

[37] D.J. Meyer and J. Meyer. *Measuring Risk Aversion*. NOW Publishers, 101 pages, 2006.

[38] R. Miculescu. Approximation of continuous functions by lipschitz functions. *Real Analysis Exchange*, 26:449–452, 2000.

[39] S. Modica and M. Scarsini. A note on comparative downside risk aversion. *Journal of Economic Theory*, 122:267–271, 2005.

[40] K.J. Murphy. *Executive Compensation*. In: Handbook of Labor Economics (O. Ashenfelter and D. Card, eds.), vol. 3b, 1999.

[41] J. Poblete and D. Spulber. The form of incentive contracts : Agency with moral hazard, risk neutrality and limited liability. *The RAND Journal of Economics*, 43:215–234, 2012.

[42] C. Prendergast. The tenuous trade-off between risk and incentives. *Journal of Political Economy*, 110(5):1071–1102, 2002.

[43] Z. Páles. On approximately convex functions. *Proceedings of the American Mathematical Society*, 131:243–252, 2003.

[44] J.F. Reuer and M.J. Leiblein. Downside risk implications of multinationality and international joint ventures. *Academy of Management Journal*, 43(2):203–214, 2000.

[45] D. Sappington. Limited liability contracts between principal and agent. *Journal of Economic Theory*, 295:1–21, 1983.

[46] B. Sinclair-Desgagné. The first-order approach to multi-signal principal-agent problems. *Econometrica*, 62:459–465, 1994.

[47] W. Whitt. Uniform conditional stochastic order. *Journal of Applied Probability*, 17:112–123, 1980.

[48] J. Wiener, M. Rogers, J. Hammitt, and P. Sand (eds). *The Reality of Precaution - Comparing Risk Regulation in the US and Europe, Resource for the Future Press*. Resource for the Future Press, Washington, 2011.

[49] O.E. Williamson. Corporate boards of directors: In principle and in practice. *Journal of Law, Economics and Organization*, 24:247–272, 2007.