

On Safeguards and Incentives*

Renaud Bourlès[†] Dominique Henriët[‡] Bernard Sinclair-Desgagné[§]

Abstract: Industrial, business and financial catastrophes often have human causes. In order to prevent such outcomes from occurring, governments, business organizations and individuals a priori invest in risk management rules, processes, technologies, and other safeguards aiming to alleviate the effects of human failures. This paper investigates how these investments combine with the incentive systems set afterwards. We find that, in both the first and the second-best, the agent will be fully insured on the downside: her wealth stays the same ex post, whether a bad or a very bad outcome realizes. We also show that the principal's safeguards investment and the agent's effort are strategic substitutes. These conclusions are subject to change, however, depending on current regulatory constraints (such as limited liability, compensation caps or due diligence requirements) and the agent's endowed wealth.

Keywords: Principal-agent analysis, risk management, tail-risk

JEL Classification: D86, L51, M12

* This paper was partly written while Sinclair-Desgagné was on sabbatical leave at GREQAM/Aix-Marseille School of Economics. We wish to thank Nathalie de Marcellis, Guy Meunier and Mihkel Tombak for helpful discussions.

[†] **Bourlès:** Ecole Centrale Marseille & GREQAM/Aix-Marseille School of Economics, rue Frédéric Joliot-Curie, 13451 MARSEILLE; renaud.bourles@centrale-marseille.fr .

[‡] **Henriët:** Ecole Centrale Marseille & GREQAM/Aix-Marseille School of Economics, rue Frédéric Joliot-Curie, 13451 MARSEILLE; dominique.henriet@centrale-marseille.fr .

[§] **Sinclair-Desgagné:** HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montréal (Québec), Canada H3T 2A7; bsd@hec.ca

1. Introduction

In the presence of significant tail risks and threats of catastrophes, cost-bearing entities typically invest in safeguards that will reduce the likelihood of disaster. Means of this sort include sensors, cameras and other warning systems installed at industrial plants, buildings or homes to alert people about a fire or the leakage of some noxious substance, double hulls equipping oil tankers in order to prevent spills, preventive bank rescue plans aiming to stabilize financial markets and avoid an economic meltdown, constitutional checks and balances meant to overcome a political body's excesses or failures, and the calls for attention our brain will automatically send us if we are about to fail dangerously on a routine task. When incentives are set in organizations or public policy, such safeguards often exist already. This paper's objectives are to examine (i) whether and how this matters for incentive provision, (ii) what the upshot might be for the probability of disaster, and (iii) how certain regulatory constraints (such as limited liability, caps on compensation, and due diligence requirements) matter in this context.

We consider these issues using a static principal-agent model baring mostly standard features (risk-averse agent, risk-neutral principal, unobservable agent effort) and three specific traits: (i) the agent's effort determines the probability of success or failure, (ii) failure may lead to bad or very bad outcomes, (iii) the principal can invest in observable safeguards which decrease the probability a failure results in disaster. This framework departs from studies of dual moral hazard (e.g., Kim and Wang 1998), in which neither the agent's nor the principal's effort are observable, or team work (e.g., Itoh 1993), in

which both the agent and the principal can monitor each other's effort perfectly. It is also distinct from works that seek to compare control systems (e.g., Kim 1995), since the safeguards we focus on here do not make the principal better or less informed. An additional peculiarity of our model is to endow the agent with certain assets whose value might be affected by the outcome; this allows to examine cases where the agent's stakes on the downside are either similar or opposed to those of the principal

In both the first and the second-best, we first find that the agent will be fully insured on the downside: her wealth stays the same ex post, whether a bad or a very bad outcome realizes. This provides a rationale for bailing out financial institutions that generate systemic risk. This conclusion may change, however, if some regulatory (or social) constraint imposes a cap on the agent's compensation in the event of a catastrophe. We also find that the principal's safeguards investment and the agent's effort are strategic substitutes. This corroborates the current literature (e.g., Daniélsson et al. 2002, Zardkoohi et al. 2016).¹ The strength of this relationship, though, and the ensuing probability of a disaster, vary according to regulation and the agent's endowment.

The rest of the paper is organized as follows. The following section introduces our mathematical notations and develops the basic model. Section 3 shows, presents and discusses our main results. The effect of widespread regulatory constraints such as limited liability, compensation caps and due diligence are considered in Sections 4, 5 and 6 respectively. Section 7 contains our concluding remarks.

¹ A notable exception is Schwartz (2016).

2. The model

Consider a principal (P) - say a regulator, a business firm, a representative voter, or the reflective part of the brain² - who must rely on an agent (A) - say, respectively, a financial institution, an executive or employee, an elected official, or the impulsive part of the brain - to achieve a good state S_1 . Realizing S_1 with probability p costs the agent $\phi(p)$. Missing S_1 , however, leads to bad or very bad states S_2 and S_3 with respective probabilities $(1-p)q$ and $(1-p)(1-q)$, where $q \in [0, 1]$ captures the effectiveness of the risk management devices, or safeguards, the principal previously implemented at cost $\psi(q)$.

Let w_1 , w_2 and w_3 represent transfers from the principal to the agent in states S_1 , S_2 and S_3 respectively. The value functions write

$$V(p, q; (w_i)) = p(v_1 - w_1) + (1-p)q(v_2 - w_2) + (1-p)(1-q)(v_3 - w_3) - \psi(q)$$

for the principal, and

$$U(p, q; (w_i)) = pu(\theta_1 + w_1) + (1-p)qu(\theta_2 + w_2) + (1-p)(1-q)u(\theta_3 + w_3) - \phi(p)$$

for the agent, where v_i (resp. θ_i) denotes the value of the principal's (resp. the agent's) assets in state S_i and $u : D \rightarrow \mathbb{R}$ is the agent's utility function.

Thereafter, we have $v_1 > v_2 > v_3$. We also take D as an interval of real numbers, and assume that u is increasing, concave and twice continuously differentiable. The cost functions ψ and φ , furthermore, are supposed to be strictly increasing, convex and twice continuously differentiable on the interval $[0, 1]$; their first derivatives are such that $\psi'(0) = \varphi'(0) = 0$ and $\psi'(q), \varphi'(p) \rightarrow \infty$ as $q, p \rightarrow 1$. Altogether, these assumptions ensure that optimal values for q and p exist, are unique, and lie strictly between 0 and 1.

² For the arguments drawn from neurobiology which support a principal-agent representation of the brain, see Brocas and Carrillo (2008).

3. Main results

Let us now first examine the first-best allocation. We shall next compare it with the second-best one.

3.1. The first-best

At the first-best, the principal will set probabilities and transfers that maximize V while insuring that the agent is willing to participate. This boils down to solving the following problem:

$$\begin{aligned} \max_{p,q,(w_i)} \quad & V(p, q; (w_i)) \\ \text{s.t.} \quad & U(p, q; (w_i)) \geq 0 \end{aligned}$$

The corresponding Lagrangean function is given by

$$\mathcal{L}^{FB}(p, q; (w_i)) = V(p, q; (w_i)) + \gamma U(p, q; (w_i))$$

where γ is the (non-negative) multiplier associated with the constraint. The first-order conditions for the optimal transfers w_1^* , w_2^* and w_3^* are then respectively given by³

$$\frac{\partial \mathcal{L}^{FB}}{\partial w_1} = -p + \gamma p u'(\theta_1 + w_1) = 0 \quad \text{so} \quad u'(\theta_1 + w_1) = \frac{1}{\gamma}, \quad (1)$$

$$\frac{\partial \mathcal{L}^{FB}}{\partial w_2} = -(1-p)q + \gamma(1-p)q u'(\theta_2 + w_2) = 0 \quad \text{so} \quad u'(\theta_2 + w_2) = \frac{1}{\gamma}, \quad (2)$$

$$\frac{\partial \mathcal{L}^{FB}}{\partial w_3} = -(1-p)(1-q) + \gamma(1-p)(1-q)u'(\theta_3 + w_3) = 0 \quad \text{so} \quad u'(\theta_3 + w_3) = \frac{1}{\gamma}. \quad (3)$$

It follows that

$$\theta_1 + w_1^* = \theta_2 + w_2^* = \theta_3 + w_3^*; \quad (4)$$

the agent thus receives the same payoff in every states and is thereby fully insured.⁴

³ Provided, of course, that $0 < p^*, q^* < 1$, and that $\gamma > 0$ (so $U(p, q; (w_i)) = 0$).

⁴ Note that this outcome does not depend on the agent's being risk averse.

Using (4), the first-order conditions for the optimal probabilities p^* and q^* are now respectively

$$\frac{\partial \mathcal{L}^{FB}}{\partial p} = (v_1 - w_1) - [q(v_2 - w_2) + (1 - q)(v_3 - w_3)] + \gamma \{u(\theta_1 + w_1) - [qu(\theta_2 + w_2) + (1 - q)u(\theta_3 + w_3)]\} - \phi'(p) = 0$$

which reduces to

$$(v_1 - w_1) - [q(v_2 - w_2) + (1 - q)(v_3 - w_3)] = \phi'(p) \quad , \text{ and} \quad (5)$$

$$\frac{\partial \mathcal{L}^{FB}}{\partial q} = (1 - p) \{[(v_2 - w_2) - (v_3 - w_3)] + \gamma[u(\theta_2 + w_2) - u(\theta_3 + w_3)]\} - \psi'(q) = 0$$

which reduces to

$$(1 - p)[(v_2 - w_2) - (v_3 - w_3)] = \psi'(q) \quad . \quad (6)$$

According to (5), the optimal probability p^* is then such that the marginal cost for the agent of delivering this likelihood of success equals the marginal benefit *to the principal*. The agent thus completely internalizes the principal's preferences. These observations are summarized in the following statement.

Proposition 1. *In the first best, the risk-averse agent is fully insured by the risk-neutral principal, i.e. her wealth stays the same whatever the outcome. The agent sets her effort p so that her marginal cost equals the principal's marginal benefit.*

We shall comment further on this, and on expression (6), in the next subsection.

3.2. The second-best

At the second-best, the principal adopts safeguards of effectiveness q^\bullet and offers the agent contingent transfers w_1^\bullet , w_2^\bullet and w_3^\bullet which maximize V , acknowledging that the agent will thereafter set a probability of success p^\bullet that maximizes U and accept this contract only

if the resulting expected value is above 0. This amounts to solving the following problem:

$$\begin{aligned} \max_{q, (w_i)} \quad & V(p^\bullet, q; (w_i)) \\ \text{s.t.} \quad & p^\bullet = \arg \max_p U(p, q; (w_i)) \\ & U(p^\bullet, q; (w_i)) \geq 0 \end{aligned}$$

The above assumptions allow to replace the first constraint - the incentive constraint - by the first-order condition for p^\bullet , which is

$$u(\theta_1 + w_1) - [qu(\theta_2 + w_2) + (1 - q)u(\theta_3 + w_3)] = \phi'(p) . \quad (7)$$

The principal's problem is then equivalent to the relaxed program

$$\begin{aligned} \max_{p, q, (w_i)} \quad & V(p, q; (w_i)) \\ \text{s.t.} \quad & u(\theta_1 + w_1) - [qu(\theta_2 + w_2) + (1 - q)u(\theta_3 + w_3)] = \phi'(p) \\ & U(p, q; (w_i)) \geq 0 \end{aligned} \quad (8)$$

The Lagrangean associated with problem (8) is given by

$$\mathcal{L}^{SB}(p, q; (w_i)) = V(p, q; (w_i)) + \lambda \{u(\theta_1 + w_1) - [qu(\theta_2 + w_2) + (1 - q)u(\theta_3 + w_3)] - \phi'(p)\} + \mu U(p, q; (w_i))$$

where λ and μ are the (non-negative) multipliers associated with the first (so-called 'incentive-compatibility') and the second (so-called 'participation') constraint respectively.

Transfers w_1^\bullet, w_2^\bullet and w_3^\bullet which solve (8) must satisfy the first-order conditions

$$\begin{aligned} \frac{\partial \mathcal{L}^{SB}}{\partial w_1} &= -p + \lambda u'(\theta_1 + w_1) + \mu p u'(\theta_1 + w_1) = 0 \\ \frac{\partial \mathcal{L}^{SB}}{\partial w_2} &= -(1 - p)q - \lambda q u'(\theta_2 + w_2) + \mu(1 - p)q u'(\theta_2 + w_2) = 0 \\ \frac{\partial \mathcal{L}^{SB}}{\partial w_3} &= -(1 - p)(1 - q) - \lambda(1 - q)u'(\theta_3 + w_3) + \mu(1 - p)(1 - q)u'(\theta_3 + w_3) = 0 \end{aligned}$$

These equations are respectively equivalent to

$$\frac{1}{u'(\theta_1 + w_1)} = \mu + \frac{\lambda}{p} \quad (9)$$

$$\frac{1}{u'(\theta_2 + w_2)} = \mu - \frac{\lambda}{1-p} \quad (10)$$

$$\frac{1}{u'(\theta_3 + w_3)} = \mu - \frac{\lambda}{1-p} \quad (11)$$

Since u is concave and the multipliers λ and μ are non-negative, we then have that

$$\theta_1 + w_1^\bullet > \theta_2 + w_2^\bullet = \theta_3 + w_3^\bullet . \quad (12)$$

The principal will thus fully insure the agent against lower-tail risk. This was also the case in the first-best, but the agent now rather gets a premium if the good state is reached. This result does not depend on whether or not the agent is downside risk-averse, or prudent.⁵

Proposition 2. *In the second best, the risk-averse agent gets a premium if the best outcome is achieved. She is also fully insured by the risk-neutral principal, i.e. her wealth stays the same, across all lower-tailed (either bad or very bad) results.*

Note that the wages themselves need not increase with the principal's valuation of states. For instance, if the agent suffers as well in state S_3 , i.e. $\theta_2 > \theta_3$, then meeting the equality in (12) implies that $w_2^\bullet < w_3^\bullet$.

Taking stock of the relationships in (12), furthermore, the agent's incentive constraint (7) becomes

$$u(\theta_1 + w_1^\bullet) - u(\theta_2 + w_2^\bullet) = \phi'(p^\bullet) . \quad (13)$$

⁵ Formally, someone is prudent when her marginal utility function is strictly convex (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). Equivalently, she prefers additional volatility to be associated with good rather than bad outcomes (Eeckhoudt and Schlesinger 2006). Note that prudence is strictly finer than risk aversion: as Crainich et al. (2013) show, risk lovers can be prudent while risk avoiders may not be.

Observe that the principal's safeguards are not explicitly present here; they remain implicit, though, through the multipliers λ and μ which serve to determine the optimal transfers.

Comparing (5) and (13), one can see that, if the difference $v_1 - v_2$ is big enough, then $p^\bullet < p^*$, so the agent's second-best effort is lower than it would be in the first-best.

Turning to the principal's optimal choice of safeguards q^\bullet , under program (8) it will satisfy the first-order condition

$$\frac{\partial \mathcal{L}^{SB}}{\partial q} = (1-p)[(v_2-w_2)-(v_3-w_3)] - \lambda[u(\theta_2+w_2)-u(\theta_3+w_3)] + \mu(1-p)[u(\theta_2+w_2)-u(\theta_3+w_3)] - \psi'(q) = 0.$$

This, by (12), comes down to

$$(1-p^\bullet)[(v_2-w_2^\bullet) - (v_3-w_3^\bullet)] = \psi'(q^\bullet). \quad (14)$$

In the benchmark case where the agent does not have any assets, i.e. $\theta_1 = \theta_2 = \theta_3 = 0$, we have that $w_1^* = w_2^* = w_3^*$ and $w_2^\bullet = w_3^\bullet$. Comparing expressions (6) and (14) with $p^\bullet < p^*$ then implies that $q^* < q^\bullet$. The principal would therefore make greater investments in safeguards in the second-best, thereby compensating for the agent's subsequent selection of a lower probability of success.

Proposition 3. *In the second best, the principal's optimal choice of safeguards q^\bullet and the agent's selected effort level p^\bullet are strategic substitutes: the principal's greater (lower) investments in safeguards has the agent deliver a lower (greater) probability of success.*

How strongly p^\bullet and q^\bullet can move in opposite directions, however, depends on the relative magnitude of θ_2 and θ_3 . Using (12), the relationship in (14) becomes

$$(1-p^\bullet)[(v_2+\theta_2) - (v_3+\theta_3)] = \psi'(q^\bullet). \quad (15)$$

If $\theta_2 < \theta_3$, for instance (so the agent's stakes on the downside are the opposite of the principal's), the factor $[(v_2 + \theta_2) - (v_3 + \theta_3)]$ is then smaller, so p^\bullet will adjust more following a change in q^\bullet . The opposite conclusion holds of course if $\theta_2 \geq \theta_3$ (in which case the principal and the agent face similar downside stakes).

The next three sections will now examine how certain regulatory constraints might affect the present conclusions.

4. Limited liability

Having in mind the case of a regulator and a bank, or that of a firm and an executive or employee, let the law impose a lower bound on compensation (so the penalties that can be inflicted to the agent). Without loss of generality, the principal's problem, as expressed in (8), must now include the additional constraints

$$w_1 \geq 0 \tag{16}$$

$$w_2 \geq 0 \tag{17}$$

$$w_3 \geq 0 \tag{18}$$

The new Lagrangean function associated with this problem is given by

$$\begin{aligned} \mathcal{L}^{LL}(p, q; (w_i)) &= V(p, q; (w_i)) + \lambda^{LL} \{u(\theta_1 + w_1) - [qu(\theta_2 + w_2) + (1 - q)u(\theta_3 + w_3)] - \phi'(p)\} \\ &\quad + \mu^{LL} U(p, q; (w_i)) + \eta_1 w_1 + \eta_2 w_2 + \eta_3 w_3 \end{aligned}$$

where η_1, η_2, η_3 are the (non-negative) multipliers associated with constraints (16), (17) and (18) respectively.

Optimal compensation will satisfy the first-order conditions

$$\begin{aligned}\frac{\partial \mathcal{L}^{SB}}{\partial w_1} &= -p + \lambda^{LL} u'(\theta_1 + w_1) + \mu^{LL} p u'(\theta_1 + w_1) + \eta_1 = 0 \\ \frac{\partial \mathcal{L}^{SB}}{\partial w_2} &= -(1-p)q - \lambda^{LL} q u'(\theta_2 + w_2) + \mu^{LL} (1-p) q u'(\theta_2 + w_2) + \eta_2 = 0 \\ \frac{\partial \mathcal{L}^{SB}}{\partial w_3} &= -(1-p)(1-q) - \lambda^{LL} (1-q) u'(\theta_3 + w_3) + \mu^{LL} (1-p)(1-q) u'(\theta_3 + w_3) + \eta_3 = 0\end{aligned}$$

These equations are respectively equivalent to

$$\frac{p - \eta_1}{p u'(\theta_1 + w_1)} = \mu^{LL} + \frac{\lambda^{LL}}{p} \quad (19)$$

$$\frac{(1-p)q - \eta_2}{(1-p)q u'(\theta_2 + w_2)} = \mu^{LL} - \frac{\lambda^{LL}}{1-p} \quad (20)$$

$$\frac{(1-p)(1-q) - \eta_3}{(1-p)(1-q) u'(\theta_3 + w_3)} = \mu^{LL} - \frac{\lambda^{LL}}{1-p} \quad (21)$$

Since the agent's incentive constraints (7) remains the same here, we may suppose that $w_1^{LL} > w_2^{LL}$ and w_3^{LL} . Hence $\eta_1 = 0$, and (19) reduces to

$$\frac{1}{u'(\theta_1 + w_1)} = \mu^{LL} + \frac{\lambda^{LL}}{p}. \quad (22)$$

In view of (20) and (21), though, the agent may not be fully insured anymore against downside contingencies if $\eta_2 > 0$ (so $w_2 = 0$) or $\eta_3 > 0$ (so $w_3 = 0$).

5. Caps on compensation

It might happen that the regulator or the firm (bending perhaps to political pressure), or even the brain (subject to negative emotions such as remorse or shame), would set a cap on the amount the agent can receive if a catastrophe happens. The following constraint is then added to the principal's problem, as expressed in (8):

$$w_3 \leq \bar{w} \quad (23)$$

The Lagrangean function now writes

$$\begin{aligned}\mathcal{L}^{cap}(p, q; (w_i)) &= V(p, q; (w_i)) + \lambda^{cap}\{u(\theta_1 + w_1) - [qu(\theta_2 + w_2) + (1 - q)u(\theta_3 + w_3)] - \phi'(p)\} \\ &\quad + \mu^{cap} \cdot U(p, q; (w_i)) + \eta \cdot (\bar{w} - w_3)\end{aligned}$$

where η is the (non-negative) multiplier associated with constraint (23).

The first-order conditions on optimal incentive pay are given by

$$\begin{aligned}\frac{\partial \mathcal{L}^{cap}}{\partial w_1} &= -p + \lambda^{cap}u'(\theta_1 + w_1) + \mu^{cap}pu'(\theta_1 + w_1) = 0 \\ \frac{\partial \mathcal{L}^{cap}}{\partial w_2} &= -(1 - p)q - \lambda^{cap}qu'(\theta_2 + w_2) + \mu^{cap}(1 - p)qu'(\theta_2 + w_2) = 0 \\ \frac{\partial \mathcal{L}^{cap}}{\partial w_3} &= -(1 - p)(1 - q) - \lambda^{cap}(1 - q)u'(\theta_3 + w_3) + \mu^{cap}(1 - p)(1 - q)u'(\theta_3 + w_3) - \eta = 0\end{aligned}$$

These equations respectively reduce to

$$\frac{1}{u'(\theta_1 + w_1)} = \mu^{cap} + \frac{\lambda^{cap}}{p} \quad (24)$$

$$\frac{1}{u'(\theta_2 + w_2)} = \mu^{cap} - \frac{\lambda^{cap}}{1 - p} \quad (25)$$

$$\frac{(1 - p)(1 - q) + \eta}{(1 - p)(1 - q)u'(\theta_3 + w_3)} = \mu^{cap} - \frac{\lambda^{cap}}{1 - p} \quad (26)$$

The upshot is that $\theta_1 + w_1^{cap} > \theta_2 + w_2^\bullet$, but the equality $\theta_2 + w_2^\bullet = \theta_3 + w_3^\bullet$ is not guaranteed if $\eta > 0$ (so $w = \bar{w}$).

6. Due diligence

Assume, finally, that (in response to the law or as required by an insurance contract, in the case of a firm; as mandated by the constitution, in a political context; or as a personal trait or developed through culture and training, in the case of the brain) the principal's safeguards are subject to some due diligence standards. The constraint

$$q \geq \bar{q} \quad (27)$$

is then added to the principal's problem expressed in (8).

The Lagrangean function is then

$$\begin{aligned}\mathcal{L}^{DD}(p, q; (w_i)) &= V(p, q; (w_i)) + \lambda^{DD}\{u(\theta_1 + w_1) - [qu(\theta_2 + w_2) + (1 - q)u(\theta_3 + w_3)] - \phi'(p)\} \\ &\quad + \mu^{DD} \cdot U(p, q; (w_i)) + \xi \cdot (q - \bar{q})\end{aligned}$$

with ξ the (non-negative) multiplier of constraint (27).

Optimal pay will meet the conditions

$$\begin{aligned}\frac{\partial \mathcal{L}^{DD}}{\partial w_1} &= -p + \lambda^{DD}u'(\theta_1 + w_1) + \mu^{DD}pu'(\theta_1 + w_1) = 0 \\ \frac{\partial \mathcal{L}^{DD}}{\partial w_2} &= -(1 - p)q - \lambda^{DD}qu'(\theta_2 + w_2) + \mu^{DD}(1 - p)qu'(\theta_2 + w_2) = 0 \\ \frac{\partial \mathcal{L}^{cap}}{\partial w_3} &= -(1 - p)(1 - q) - \lambda^{DD}(1 - q)u'(\theta_3 + w_3) + \mu^{DD}(1 - p)(1 - q)u'(\theta_3 + w_3) = 0\end{aligned}$$

These equations respectively reduce to

$$\frac{1}{u'(\theta_1 + w_1)} = \mu^{DD} + \frac{\lambda^{DD}}{p} \quad (28)$$

$$\frac{1}{u'(\theta_2 + w_2)} = \mu^{DD} - \frac{\lambda^{DD}}{1 - p} \quad (29)$$

$$\frac{1}{u'(\theta_3 + w_3)} = \mu^{DD} - \frac{\lambda^{DD}}{1 - p} \quad (30)$$

The relationship

$$\theta_1 + w_1^{DD} > \theta_2 + w_2^{DD} = \theta_3 + w_3^{DD}$$

thus holds, as in the second-best.

On the other hand, the principal's optimal safeguards q^{DD} must meet the first-order condition

$$\begin{aligned}0 &= \frac{\partial \mathcal{L}^{SB}}{\partial q} = (1 - p)[(v_2 - w_2) - (v_3 - w_3)] - \lambda[u(\theta_2 + w_2) - u(\theta_3 + w_3)] \\ &\quad + \mu(1 - p)[u(\theta_2 + w_2) - u(\theta_3 + w_3)] - \psi'(q) + \xi \ .\end{aligned}$$

which comes down to

$$(1 - p^{DD})[(v_2 + \theta_2) - (v_3 + \theta_3)] + \xi = \psi'(q^{DD}) . \quad (32)$$

If $\xi > 0$ so the lower bound on safeguards is hit ($q^{DD} = \bar{q}$), it is then possible for p^{DD} to go up without any change in q^{DD} . The principal's safeguards and the agent's effort are thus not any more strategic substitutes.

7. Conclusion

This paper investigated the relationship between incentive provision and the presence of safeguards to mitigate downside risk. We found that, in both the first and the second best, the agent should be fully insured against downside risk. Also, the agent's effort to lower the probability of failure and the principal's safeguards investment to alleviate the impact of failure are strategic substitutes. These results must be qualified, however, in the presence of limited liability, compensation caps or due diligence requirements.

One next step at this stage would be to modify the agent's utility function according to certain observations from behavioral economics or the economics of training. For instance, one might think that the agent's intrinsic motivation could be affected by the amount of safeguards (a high amount would affect her morale, for it would be seen as lack of confidence on the part of the principal). On the other hand, safeguards might consist in investments in ergonomics which will decrease the agent's marginal cost of effort. Intuitively, the former will tend to reinforce safeguards as strategic substitutes for the agent's effort, while the latter would have the opposite effect. In both cases, the resulting probability of catastrophe remains to be investigated.

References

- Brocas, Isabelle, and Juan D. Carrillo. 2008. "The Brain as a Hierarchical Organization," *American Economic Review* 98(4), p. 1312-1346.
- Crainich, David, Louis Eeckhoudt, and Alain Trannoy. 2013. "Even (Mixed) Risk Lovers are Prudent," *American Economic Review* 103(4), p. 1529-1535.
- Danielsson, Jón, Bjørn N. Jorgensen, and Casper G. de Vries. 2002. "Incentives in Effective Risk Management," *Journal of Banking and Finance* 26, p. 1407-1425.
- Eeckhoudt, Louis, and Harris Schlesinger. 2006. "Putting Risk in its Proper Place," *American Economic Review* 96, p. 280-289.
- Itoh, Hideshi. 1993. "Coalitions, Incentives, and Risk Sharing," *Journal of Economic Theory* 60(2), p. 410-427.
- Kim, Son Ku. 1995. "Efficiency of an Information System in an Agency Model," *Econometrica* 63(1), p. 89-102.
- Kim, Son Ku, and Susheng Wang. 1998. "Linear Contracts and the Double Moral-Hazard," *Journal of Economic Theory* 82(2) p. 342-378.
- Kimball, Miles. 1990. "Precautionary Saving in the Small and in the Large," *Econometrica* 5, p. :53-73.
- Menezes, Carmen, C. Geiss, and J. Tressler. 1980. "Increasing Downside Risk," *American Economic Review* 70, p. :921-932.
- Schwartz, Steven L. 2017. "Too Big Too Fool: Moral Hazard, Bailouts, and Corporate Responsibility," Duke Law School Public Law & Legal Theory Series No. 2016-58
- Zardkoochi, Asghar, Eugene Kang, Donald Fraser, and Albert A. Cannella. 2016. "Managerial Risk-Taking Behavior: A Too-Big-To Fail Story," *Journal of Business Ethics*, published online 31 March 2016.