On Safeguards and Incentives*

Renaud Bourlès† Dominique Henriet‡ Bernard Sinclair-Desgagné§

August 2017

Preliminary and incomplete

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, Mihkel Tombak and participants to a seminar at Aix-Marseille School of Economics for helpful discussions.

†Centrale Marseille (Aix-Marseille School of Economics), rue Frédéric Joliot-Curie, 13451 MARSEILLE; renaud.bourles@centrale-marseille.fr .

‡Centrale Marseille (Aix-Marseille School of Economics), rue Frédéric Joliot-Curie, 13451 MARSEILLE; dominique.henriet@centrale-marseille.fr .

§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 et 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. Daníelsson *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.

2 The model

Consider a principal (P) - say a regulator, a business firm, a representative voter, or the reflective part of the brain[2] - 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 probabilies $(1-p)q$ 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

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

[1] A notable exception is Schwarcz (2017).

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

for the principal, and

$$U(p,q;(w_i)) = pu(\theta_1 + w_1) + (1-p)qu(\theta_2 + w_2) + (1-p)(1-q)u(\theta_3 + w_3) - \phi(p) \quad (2)$$

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 (with $s_i \equiv v_i + \theta_i$) and $u : D \to \mathbb{R}$ is the agent's utility function.

Thereafter, we assume $v_1 > v_2 > v_3$ and $s_1 > s_2 > s_3$. 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 $[0, 1]$.

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:

$$\begin{aligned} \max_{p,q,(w_i)} \quad & V(p,q;(w_i)) \\ \text{s.t.} \quad & U(p,q;(w_i)) \geq 0 \end{aligned} \quad (3)$$

We focus in the following on the internal solution.

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:*

$$\theta_1 + w_1^* = \theta_2 + w_2^* = \theta_3 + w_3^* = u^{-1}(\phi(p^*)) \quad (4)$$

Agent's effort p^* and principal safeguards q^* are set so that their marginal cost equal their social marginal benefit:

$$\begin{cases} \phi'(p^*).\left[u^{-1}(\phi(p^*))\right]' & = & s_1 - q^* s_2 - (1-q^*)s_3 \\ \psi'(q^*) & = & (1-p^*)(s_2 - s_3) \end{cases} \tag{5}$$

Proof*: See Appendix 8.1.*

We shall comment further on this 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:

$$\begin{aligned} \max_{q,(w_i)} \quad & V(p^\bullet, q; (w_i)) \\ \text{s.t.} \quad & p^\bullet = \arg\max_p U(p, q; (w_i)) \\ & U(p^\bullet, q; (w_i)) \geq 0 \end{aligned} \tag{6}$$

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

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

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

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

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.*

$$\theta_1 + w_1^\bullet > \theta_2 + w_2^\bullet = \theta_3 + w_3^\bullet \tag{9}$$

More precisely:

$$\begin{cases} \theta_1 + w_1^\bullet & = \quad u^{-1}(\phi(p^\bullet) + (1-p^\bullet)\phi'(p^\bullet)) \\ \theta_2 + w_2^\bullet = \theta_3 + w_3^\bullet & = \quad u^{-1}(\phi(p^\bullet) - p^\bullet\phi'(p^\bullet)) \end{cases} \tag{10}$$

Proof*: See Appendix 8.2.*

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.[3]

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. $\theta_2 > \theta_3$, as argued by Schwarcz (2017) in the case of financial crisis, meeting the equality in (9) implies that $w_2^\bullet < w_3^\bullet$. Regulation as caps on compensation might then distort the optimum (see section 5).

[3] 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 et 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.

Still, assuming $\phi(0) = 0$, the second equation of (10) would entail $\theta_2 + w_2^\bullet = \theta_3 + w_3^\bullet < 0$ as usual in a moral hazard context.[4] Liability regulation is therefore also likely to play an important role (see section 4).

Now turn to the principal's optimal choice of safeguards $q^\bullet$.

Proposition 3 *In the second best, as in the first best:*

$$\psi'(q^\bullet) = (1 - p^\bullet)(s_2 - s_3) \tag{11}$$

Therefore, as $s_2 > s_3$, 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.*
Proof*: See Appendix 8.3.*

How strongly $p^\bullet$ and $q^\bullet$ can move in opposite directions, however, depends on the relative magnitude of s_2 and s_3. If $\theta_2 < \theta_3$, for instance (so the agent's stakes on the downside are the opposite of the principal's), the factor $s_2 - s_3$ is then smaller (recall here that $s_i = v_i + \theta_i$), 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).

Let us now compare the levels of effort and safeguards of the first and the second best. Using the properties of the optimal transfers (equation (4) and (10)) we can rewrite the program of the first and second best respectively as:

$$(p^*, q^*) = \arg\max_{p,q} W_{FB} \equiv \left[ps_1 + (1-p)qs_2 + (1-p)(1-q)s_3 - \psi(q) - u^{-1}(\phi(p)) \right] \tag{12}$$

and

$$\begin{aligned} (p^\bullet, q^\bullet) = \arg\max_{p,q} W_{SB} \quad &\equiv \quad \Big[ps_1 + (1-p)qs_2 + (1-p)(1-q)s_3 - \psi(q) \\ &- \quad pu^{-1}\big(\phi(p) + (1-p)\phi'(p)\big) - (1-p)u^{-1}\big(\phi(p) - p\phi'(p)\big) \Big] \end{aligned} \tag{13}$$

[4] Recall here that $\phi(.)$ is convex and that $u(0) = 0$ as can be seen from the participation constraint

Then, the objective of the the second can be written as $W_{SB} = W_{FB} - R(p)$ where $R(p) \equiv pu^{-1}(\phi(p) + (1-p)\phi'(p)) + (1-p)u^{-1}(\phi(p) - p\phi'(p)) - u^{-1}(\phi(p))$ can be understood as a risk premium for the agent. Indeed, whereas in the first first best, the agent only has to be compensated for participation (as reflected by the term $u^{-1}(\phi(p))$ in W_{FB}), the second best entails an increase in risk from the point of the agent (as then $v_1+\theta_1 \neq v_2+\theta_2 = v_3+\theta_3$), making participation of the agent more costly for the principal ($u(\cdot)$ being concave and $\phi(\cdot)$ convex, $R(p) > 0 \; \forall p \in (0,1)$). The shape of this premium then determines whether moral hazard increases effort or safeguards.

Proposition 4 *If $0 < \phi'(0) < \phi'(1) < +\infty$, then $R'(1) < 0 < R'(0)$, and $\exists \; (\underline{p}, \overline{p}) \in (0,1)^2$ with $\underline{p} \leq \overline{p}$ such that, moral hazard involves*

- *less effort ($p^\bullet < p^*$) and more safeguards ($q^\bullet > q^*$) if $p^* < \underline{p}$*
- *more effort ($p^\bullet > p^*$) and less safeguards ($q^\bullet < q^*$) if $p^* > \overline{p}$*

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 $w_i \geq 0$, $i = 1, 2, 3$. In such cases, the agent may not be fully insured anymore against downside contingencies, breaking down the result on substituability of effort and safeguards.

[IN PROGRESS]

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). Such a regulation then entails risk in the downside for the agents, modifying the link between efforts and safeguards. Caps on compensation mechanically increases transfers in states S_1 and S_2, to ensure participation; and reduces loss for the principal in state S_3. They are therefore likely to decrease safeguards.

[IN PROGRESS]

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 personality 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 $q \geq \overline{q}$ is then added to the principal's problem expressed in (8). The substituability between principal's safeguards and the agent's effort is then no longer guaranteed (when the constraint is binding).

[IN PROGRESS]

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 agains 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.

8 Appendix

8.1 Proof of Proposition 1

The Lagrangean of program (3) 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^*, optimal effort p^* and optimal safeguards q^* are then respectively given by[5]

$$\begin{aligned}
\frac{\partial \mathcal{L}^{FB}}{\partial w_1} &= -p + \gamma p u'(\theta_1 + w_1) = 0 \ , && (14)\\
\frac{\partial \mathcal{L}^{FB}}{\partial w_2} &= -(1-p)q + \gamma(1-p)qu'(\theta_2 + w_2) = 0 \ , && (15)\\
\frac{\partial \mathcal{L}^{FB}}{\partial w_3} &= -(1-p)(1-q) + \gamma(1-p)(1-q)u'(\theta_3 + w_3) = 0 \ , && (16)\\
\frac{\partial \mathcal{L}^{FB}}{\partial p} &= (v_1 - w_1) - [q(v_2 - w_2) + (1-q)(v_3 - w_3)] \\
&\quad + \gamma\{u(\theta_1 + w_1) - [qu(\theta_2 + w_2) + (1-q)u(\theta_3 + w_3)] - \phi'(p)\} = 0 \ , && (17)\\
\frac{\partial \mathcal{L}^{FB}}{\partial q} &= (1-p)\{[(v_2 - w_2) - (v_3 - w_3)] + \gamma[u(\theta_2 + w_2) - u(\theta_3 + w_3)]\} \\
&\quad - \psi'(q) = 0 \ . && (18)
\end{aligned}$$

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

It follows from the first three equations that

$$\theta_1 + w_1^* = \theta_2 + w_2^* = \theta_3 + w_3^* \tag{19}$$

and $\gamma = 1/u'(\theta_i + w_i^*)$, $i = \{1,2,3\}$. The binding participation constraints then gives $\theta_i + w_i^* = u^{-1}(\phi(p^*))$, $i = \{1,2,3\}$.

Equation (17), then gives

$$(v_1 - w_1^*) - [q^*(v_2 - w_2^*) + (1-q)(v_3 - w_3^*)] = \phi'(p^*)\gamma \tag{20}$$

That is using (19) and the definition of γ

$$s_1 - q^* s_2 - (1-q^*)s_3 = \phi'(p^*).\left[u^{-1}(\phi(p^*))\right]' \tag{21}$$

Similarly (18) gives

$$(1-p^*)(s_2 - s_3) = \psi'(q^*) \tag{22}$$

8.2 Proof of Proposition 2

The Lagrangean associated with problem (8) is given by

$$\mathcal{L}^{SB}(p,q;(w_i)) = V(p,q;(w_i)) + \lambda\{u(\theta_1+w_1) - [qu(\theta_2+w_2)+(1-q)u(\theta_3+w_3)] - \phi'(p)\} + \mu U(p,q;(w_i)) \tag{23}$$

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

$$\frac{\partial \mathcal{L}^{SB}}{\partial w_1} = -p + \lambda u'(\theta_1 + w_1) + \mu p u'(\theta_1 + w_1) = 0 \tag{24}$$

$$\frac{\partial \mathcal{L}^{SB}}{\partial w_2} = -(1-p)q - \lambda q u'(\theta_2 + w_2) + \mu(1-p)qu'(\theta_2 + w_2) = 0 \tag{25}$$

$$\frac{\partial \mathcal{L}^{SB}}{\partial w_3} = -(1-p)(1-q) - \lambda(1-q)u'(\theta_3 + w_3) + \mu(1-p)(1-q)u'(\theta_3 + w_3) = 0 \tag{26}$$

These equations are respectively equivalent to

$$\frac{1}{u'(\theta_1+w_1)} = \mu + \frac{\lambda}{p} \tag{27}$$
$$\frac{1}{u'(\theta_2+w_2)} = \mu - \frac{\lambda}{1-p} \tag{28}$$
$$\frac{1}{u'(\theta_3+w_3)} = \mu - \frac{\lambda}{1-p} \tag{29}$$

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

$$\theta_1 + w_1^\bullet > \theta_2 + w_2^\bullet = \theta_3 + w_3^\bullet\ . \tag{30}$$

The binding constraints then give

$$\begin{cases} pu(\theta_1 + w_1^\bullet) + (1-p)u(\theta_2 + w_2^\bullet) = \phi(p^\bullet) \\ u'(\theta_1 + w_1^\bullet) - u'(\theta_2 + w_2^\bullet) = \phi'(p^\bullet) \end{cases} \tag{31}$$

that is

$$\begin{cases} u(\theta_1 + w_1^\bullet) = \phi(p^\bullet) + (1-p)\phi'(p^\bullet) \\ u(\theta_2 + w_2^\bullet) = \phi(p^\bullet) - p\phi'(p^\bullet) \end{cases} \tag{32}$$

8.3 Proof of Proposition 3

The first-order condition of program (8) with respect to the level of safeguards writes

$$\frac{\partial \mathcal{L}^{SB}}{\partial q} = (1-p)[(v_2-w_2)-(v_3-w_3)]-\lambda[u(\theta_2+w_2)-u(\theta_3+w_3)]+\mu(1-p)[u(\theta_2+w_2)-u(\theta_3+w_3)]-\psi'(q) = 0. \tag{33}$$

As $\theta_2 + w_2^\bullet = \theta_3 + w_3^\bullet$, this gives

$$(1-p^\bullet)[(v_2 - w_2^\bullet) - (v_3 - w_3^\bullet)] = \psi'(q^\bullet) \tag{34}$$

that can be written as

$$(1-p^\bullet)(s_2 - s_3) = \psi'(q^\bullet) \tag{35}$$

as then, $w_2^\bullet - w_3^\bullet = \theta_2 - \theta_3$.

References

Brocas, I. et Carrillo, J. D. (2008). The brain as a hierarchical organization. *American Economic Review*, 98(4):1312–46.

Crainich, D., Eeckhoudt, L. et Trannoy, A. (2013). Even (mixed) risk lovers are prudent. *American Economic Review*, 103(4):1529–35.

Daníelsson, J., Jorgensen, B. N. et de Vries, C. G. (2002). Incentives for effective risk management. *Journal of Banking & Finance*, 26(7):1407 – 1425.

Eeckhoudt, L. et Schlesinger, H. (2006). Putting risk in its proper place. *American Economic Review*, 96(1):280–289.

Itoh, H. (1993). Coalitions, incentives, and risk sharing. *Journal of Economic Theory*, 60(2):410 – 427.

Kim, S. K. (1995). Efficiency of an information system in an agency model. *Econometrica*, 63(1):89–102.

Kim, S. K. et Wang, S. (1998). Linear contracts and the double moral-hazard. *Journal of Economic Theory*, 82(2):342 – 378.

Kimball, M. S. (1990). Precautionary Saving in the Small and in the Large. *Econometrica*, 58(1):53–73.

Menezes, C., Geiss, C. et Tressler, J. (1980). Increasing downside risk. *American Economic Review*, 70(5):921–932.

Schwarcz, S. L. (2017). Too big to fool: Moral hazard, bailouts, and corporate responsibility. *Minnesota Law Review*, forthcoming.

Zardkoohi, A., Kang, E., Fraser, D. et Cannella, A. A. (2016). Managerial risk-taking behavior: A too-big-to-fail story. *Journal of Business Ethics.*