games

MDPI

Article

The Optimal Contract under Adverse Selection in a Moral-Hazard Model with a Risk-Averse Agent

François Maréchal and Lionel Thomas *

CRESE EA3190, University Bourgogne Franche-Comté, F-25000 Besançon, France; francois.marechal@univ-fcomte.fr

* Correspondence: lionel.thomas@univ-fcomte.fr

Received: 13 November 2017; Accepted: 21 February 2018; Published: 1 March 2018

Abstract: This paper studies the optimal contract offered by a risk-neutral principal to a risk-averse agent when the agent's hidden ability and action both improve the probability of the project being successful. We show that if the agent is sufficiently prudent and able, the principal induces a higher probability of success than under moral hazard, despite the costly informational rent given up. Moreover, there is distortion at the top. Finally, the conditions to avoid pooling are difficult to satisfy because of the different kinds of incentives to be managed and the overall trade-off between rent extraction, insurance, and efficiency involved.

Keywords: adverse selection; moral hazard; risk aversion; prudence

JEL Classification Number: D82

---

1. Introduction

The theory of incentives has made considerable advances in the last forty years. The implications of pure adverse-selection or pure moral-hazard models are now well known.[1] However, there are many examples of contracts designed to solve adverse-selection and moral-hazard problems simultaneously. Chief executive officer (CEO) and financial contracts are particularly archetypal. In the former, a CEO has private information on how good a manager he is and how tirelessly he works. In the latter, a borrower has private information on how risky his project is and how hazardously he carries it out. Paradoxically, despite the plethora of mixed models in the incentives literature, such problems where the principal's payoff is closely bound up with the agent's private information and unobservable action have come in for little study.

In this paper, a stochastic output taking two values, low (i.e., failure) or high (i.e., success), is considered in order to study the optimal contract offered by a risk-neutral principal to a risk-averse agent when the agent's hidden ability (or type) and hidden action both improve the probability of success.

Three key elements arise. First, if ability and prudence, relative to the contract offered, are sufficiently high, the equilibrium contract induces some agents to exert more effort than they normally would in a moral-hazard alone setup.[2] Second, the contract offered to the most able type is distorted in order to reduce the cost of providing incentives. Third, the principal may optimally

---

1 See Laffont and Martimort (2002) [1], Bolton and Dewatripont (2005) [2].

2 In this paper, we adopt the following convention. A pure moral-hazard or adverse-selection problem contains only one item of asymmetric information. A moral-hazard or adverse-selection model alone contains both action and ability, but only one of these is unobservable.

pool some types by offering a contract with constant transfers. To illustrate the intuitions behind these results, let us describe the properties of full-information and moral-hazard alone contracts.

We begin by recalling the well-known full-insurance property of the full-information contract: to generate a positive probability of success, the agent receives a (full-information) fixed payment whether the production fails or succeeds. Furthermore, he gets no rent (i.e., he only receives his reservation utility). Then, we consider that the agent's action becomes non-observable, and we recall the main characteristics of the moral-hazard contract. The agreement offers high-powered incentives: the (moral-hazard) fixed payment is increased by a positive bonus in the event of success. Moreover, because the agent bears some risk, the contract pays him a risk-premium without giving up a rent for him. These elements constitute the expected payment made to the agent or the moral-hazard cost incurred by the principal to induce a positive probability of success despite the non-observability of the action. To take this cost into account, it is optimal to distort the efficient probability of success. More precisely, the probability is increased until the marginal benefit (i.e., the increase in output value) equals the marginal expected payment (i.e., the moral-hazard marginal cost of increasing the likelihood of success). Such distortions reflect the usual trade-off between insurance and efficiency.

Next, we consider that ability is no longer observable. Thus, adverse selection complements moral hazard to form a mixed model. When the principal faces asymmetric information about the agent's type, she needs to elicit truthful information. To do this, she must give up an informational rent to the agent. Compared with moral hazard, such a rent introduces two modifications of the cost incurred by the principal. First, this rent constitutes the usual adverse-selection cost. It compensates the agent's incentives to mimic an agent with low ability. It follows that it is null, also at the margin, for the most able agent. This cost is added to the moral-hazard cost to form the mixed-model cost.

Second, the rent affects the two components of the moral-hazard cost. The consequence for the bonus is inevitably costly. Because the bonus incentivizes the risk-averse agent to bear a risk, the higher the fixed payment is, the more costly high-powered incentives are. That is precisely what the informational rent does since the agent gets no rent under moral hazard alone.

By contrast, the consequence of the informational rent for the risk premium is ambiguous. It is important to notice that a rent is equivalent to transferring the risk borne by the agent toward higher utility levels than the reservation utility. This forces the principal to pay a risk premium that guarantees a higher utility in the mixed model than in the moral-hazard model. On the one hand, it is costly at the margin because of risk aversion again. On the other hand, in accordance with decision theory, this transfer is not welcomed by an imprudent agent, but the contrary is true for a prudent one.[3] In the former case, this prompts an increase in the risk premium (at the margin). In the latter, the contrary arises.

Ultimately, the presence of adverse selection can reinforce or mitigate the moral-hazard trade-off. More specifically, the informational rent given up implies that:

- risk aversion and prudence both contribute to increasing the moral-hazard marginal cost to induce a positive probability of success if the agent is imprudent,
- prudence softens the increase in the moral-hazard marginal cost due to risk aversion if the agent is prudent, and even more than offsets the increase if prudence is sufficiently high (i.e., above an endogenous level).[4]

---

[3] The notion of prudence initially concerns decision theory. The formal definition is due to Kimball (1990) [3] and is linked to the choice of a level of precautionary saving. Since then, it has been shown that prudence can also be interpreted as a preference towards risk, as risk aversion. Menezes, Geiss and Tressler (1980) [4] study "prudence" through the notion of downside risk aversion. Eeckhoudt, Gollier and Schneider (1995) [5] show that upward shifts of any increase in risk are beneficial to expected utility if and only if the individual is prudent. From this result, it is possible to show that transferring a risk toward higher levels of wealth is appreciated by a prudent agent (see Crainich and Eeckhoudt, 2005) [6]. In our paper, we make the connection between wealth and utility levels.

[4] The notion of sufficiently highly prudent agent will be stated rigorously in the core of the paper.

It follows that the optimal probability is such that the marginal benefit equals the mixed-model marginal cost, i.e., the sum of the moral-hazard marginal cost and the adverse-selection marginal cost (i.e., the informational-rent marginal cost).

This allows us to state our first result. A surprising effect arises if the agent is both sufficiently highly prudent and sufficiently able. If prudence is sufficiently high, the informational rent left by the contract implies a reduction in the moral-hazard marginal cost. In parallel, this informational rent leads to an adverse-selection cost. However, in accordance with pure adverse-selection models, a sufficiently able agent implies an adverse-selection marginal cost relatively close to zero (even null at the top). Therefore, in the presence of both a sufficiently highly prudent and a sufficiently able agent, the decrease in the moral-hazard marginal cost more than offsets the adverse-selection marginal cost. It follows that the mixed-model contract entails a higher probability of success than the moral-hazard contract, despite the costly informational rent given up. Otherwise, there is a reduction in the probability of success when the agent is imprudent or insufficiently prudent since the informational rent also increases the moral-hazard marginal cost due to risk aversion.

This implies our second result. Since the informational rent also implies a modification in the moral-hazard marginal cost, there is inevitably a distortion for all the agent's ability compared to moral hazard alone, even at the top. Thus, this mixed model implies more distortions than a pure adverse-selection model. Such distortions reflect the overall trade-off between rent extraction, insurance, and efficiency.

Nevertheless, the nature of these distortions varies with the agent's ability. We show that moving from the lowest to the highest ability leads the overall trade-off to substitute distortions due to moral hazard for distortions due to adverse selection. Indeed, again according to adverse-selection models alone, rather low types obtain a low rent, which implies a little modification in the moral-hazard marginal cost, and are associated with a high marginal cost of the informational rent. The contrary is true for rather able agents. Therefore, the distortion associated with low types is due rather to the rent extraction-efficiency trade-off. The principal is more concerned about the adverse-selection cost. The probability of success is distorted to limit the informational rent. The reverse is true for the distortion associated with high types. It occurs from the insurance-efficiency trade-off, because the principal is more interested in the moral-hazard cost. The distortion comes from the desire to limit the cost to induce a positive probability of success, i.e., the bonus and the risk premium.

Our third result follows. Indeed, this substitution of the distortion effect raises the question whether a fully separating contract can be implemented. Pooling can occur because of the non-responsiveness phenomenon (i.e., the non-implementability of the moral-hazard contract) and the possible lack of monotonicity of the marginal cost of the informational rent due to risk aversion. Therefore, pooling in the mixed contract does not have different causes from those already known in pure adverse-selection models. However, pooling is most likely to emerge since the conditions for avoiding pooling are more difficult to satisfy because of the different kinds of incentives to be managed and the overall trade-off between rent extraction, insurance, and efficiency involved.

The paper is organized as follows. Section 2 presents the related literature. The model is stated in Section 3. Sections 4 and 5 respectively analyze full-information and moral hazard. The mixed case is studied in Section 6. An example is given in Section 7. We briefly conclude in an eighth section. Appendix A is devoted to appendices.

2. Related Literature

The literature on screening under moral-hazard can be split into three distinct parts. The first part concerns models similar to ours. Faynzilberg and Kumar (2000) [7] analyze an equivalent model, but with a continuum of output. They study more specifically the validity of the well-known first-order approach. However, this degree of generality implies that the optimal contract cannot be fully characterized. On the one hand, the two-output model constitutes a limit. However, on the other hand, it allows us to use the first-order approach and to fully determine the properties of the optimal contract and the overall trade-off between rent extraction, insurance, and efficiency.

Ollier (2007) [8] and Ollier and Thomas (2013) [9] also study a two-output model, but with a risk-neutral agent protected by limited liability or ex-post participation constraints. They show that a fully pooling contract is optimal. Such a contract allows the principal to reduce the overall (informational plus limited liability or ex-post participation) rent left to the agent. In our model, pooling can also arise. However, this is due to already known reasons: the non-responsiveness phenomenon or the lack of monotonicity of the marginal cost of the informational rent. Therefore, pooling is not a cause for the principal to limit rent, but a consequence of the difficulty in managing different kinds of incentives in the presence of risk aversion. A similar paper, except for liability protections, by Escobar and Pulgar (2017) [10] shows that separation is also a difficult task due to the basic tensions that arise because the agent could deviate in two different dimensions.

The second part of the literature studies insurance models. In Jullien et al. (2007) [11], a two-type/two-output model is analyzed. The major difference with the present paper is that the agent's private knowledge affects the level of risk-aversion instead of the technology. One result can be underlined: the properties of the optimal contract crucially depend on the high power incentives existing in the outside option. In our paper, we do not investigate an endogenous outside-option utility. By contrast, we assume that it is constant.

The third part comes from the fact that our mixed setup is a model in which true moral hazard follows adverse selection. It differs from a mixed model with "false" moral hazard. With false moral hazard, the action undertaken by the agent deterministically affects the variable observed by the principal. The agent receives an information rent only and pure adverse selection trade-offs arise (e.g., Laffont and Tirole, 1986) [12]: efficiency at the top, downward distortions otherwise and a fully-separating contract under usual regularity conditions.

3. The Model

The basic data of the mixed problem follow Ollier and Thomas (2013) [9]. A principal contracts with an agent to produce a pecuniary output with random value $x, x \in \{\underline{x}, \overline{x}\}$. High (resp. low) output $\overline{x}$ (resp. $\underline{x}$) is associated with success (resp. failure). The realization of a high output requires the agent to perform an action generating a probability of success $\rho = \Pr(x = \overline{x}) \in (0, 1)$. However, action is costly for the agent. He incurs an indirect disutility $\psi(\rho, \theta)$, with θ his ability. The principal does not observe the action or the ability. However, she knows that ability is drawn from a density $f > 0$ on $[\underline{\theta}, \overline{\theta}]$ with cumulative F. The principal offers a contract $\langle a, b \rangle$ with a, a fixed payment, and b, a bonus in the event of success. In other words, a is the non-contingent component of the contract and b the contingent component. Let U_0 be the agent's reservation utility.

The principal. The principal is risk-neutral. For a given probability of success and a given contract, her objective function is:

$$\begin{aligned} V &= (1-\rho)\big(\underline{x} - a\big) + \rho\big(\overline{x} - (a+b)\big) \\ &= \underline{x} + \rho \Delta x - a - \rho b, \end{aligned} \tag{1}$$

where $\Delta x = \overline{x} - \underline{x} > 0$ is the increase in output value due to success.

The agent. The agent is risk-averse. The utility of the payment in the event of success (resp. failure) is $u(a+b)$ (resp. $u(a)$). The agent's expected utility is:

$$\begin{aligned} U &= (1-\rho)u(a) + \rho u(a+b) - \psi(\rho, \theta) \\ &= u(a) + \rho\big(u(a+b) - u(a)\big) - \psi(\rho, \theta). \end{aligned} \tag{2}$$

The properties of the functions u and ψ are the following.

Assumption 1. *The functions u and ψ satisfy:*

$$u(0) = 0, u'(.) > 0, u''(.) < 0, u'''(.) \text{ has a constant sign,}$$

and [5]*:*

$$\psi(0,\theta) = 0, \psi_1(\rho,\theta) > 0, \psi_{11}(\rho,\theta) > 0, \psi_2(\rho,\theta) < 0, \psi_{22}(\rho,\theta) > 0, \psi_{12}(\rho,\theta) < 0.$$

The conditions on the function u imply that the agent has no utility if he receives no payment, is risk-averse since the marginal utility of the payment is positive and decreasing, and is either imprudent when $u''' < 0$ or prudent when $u''' > 0$. The conditions on the function ψ reflect usual assumptions about the probability of success and the disutility of effort.[6] For all types, the disutility incurred to implement no probability of success is null. The marginal indirect disutility is positive and increasing. In more productive states, the agent's disutility diminishes, but at an increasing rate. Higher values of θ correspond to states in which a higher probability of success is less costly to generate.

Let $r_A = -\frac{u''}{u'}$ (resp. $r_P = -\frac{u'''}{u''}$) be the absolute risk aversion (resp. prudence) coefficient. For the analysis, it is useful to introduce different levels of prudence.

Definition 1. *The agent is said to be,* $\forall(\rho,\theta) \in (0,1) \times [\underline{\theta}, \overline{\theta}]$,

- *weakly prudent if*
$$0 < r_P \leq 3r_A,$$
- *highly prudent if*
$$3r_A < r_P.$$

The timing. The timing of this contracting game unfolds as follows:

- the agents learns his type θ;
- the principal offers a contract to the agent. If the agent rejects, the game ends; if he accepts, he chooses an unobservable action;
- the principal observes the output and the contract is executed.

The problem. The principal's problem is to maximize the expectation of (1) with respect to $\langle a, b\rangle$, subject to the participation and the incentive compatibility constraints. The former implies that the agent voluntarily agrees to the contract. The latter requires the agent to be honest and obedient.

To determine the contract properties and to favor the analysis of the distortions imputable to information asymmetries, it will be useful to focus the study on the rent (if any) left to the agent. To do so, we adopt the following change of variables. Let $\underline{u}$ be the utility in the event of failure, $\overline{u}$ the utility in the event of success, and Δu the spread of utility (hereafter, the power of incentives). We obtain:

$$\begin{cases} \underline{u} = u(a), \\ \overline{u} = u(a+b), \\ \Delta u = \overline{u} - \underline{u}. \end{cases} \tag{3}$$

Denoting $u^{-1} = w$, we get the first lemma.

---

5 Subscript i denotes the partial derivative with the i-th argument.

6 Indeed, an equivalent way to analyze this problem is to consider that the agent exerts an effort that, like his ability, increases the probability of success $\rho(e,\theta)$. However, he incurs a disutility $\varphi(e)$. See Ollier and Thomas (2013) [9] for details.

Lemma 1. *The function w is such that:*

$$
\begin{aligned}
w' &= \frac{1}{u'} > 0, \\
w'' &= -\frac{u''}{u'^3} > 0, \\
w''' &= -\frac{u'''u' - 3u''^2}{u'^5} \lesseqgtr 0 \Leftrightarrow r_P \gtreqless 3r_A.
\end{aligned}
$$

Proof. Straightforward. □

The signs of w' and w'' directly follow from Assumption 1. Using Definition 1, the sign of w''' depends on the level of the agent's prudence. If the agent is imprudent or weakly prudent (resp. highly prudent), w''' is positive (resp. negative).

Considering payments, we get the following lemma.

Lemma 2. *The payments are such that:*

$$
\begin{aligned}
a &= w(\underline{u}), \\
b &= w(\overline{u}) - w(\underline{u}).
\end{aligned}
$$

Proof. Straightforward. □

Using this lemma, the objective function (1) is:

$$
V = \underline{x} + \rho\Delta x - w(\underline{u}) - \rho\big(w(\overline{u}) - w(\underline{u})\big). \tag{4}
$$

From (3), the agent's utility (2) becomes:

$$
U = \underline{u} + \rho\Delta u - \psi(\rho, \theta). \tag{5}
$$

Moreover, manipulating (5) and since $\overline{u} = \Delta u + \underline{u}$, we get:

$$
\begin{cases}
\underline{u} = U + \psi(\rho, \theta) - \rho\Delta u, \\
\bar{u} = U + \psi(\rho, \theta) + \big(1 - \rho\big)\Delta u.
\end{cases} \tag{6}
$$

Thus, offering the contract $\langle U, \Delta u\rangle$ specifying the agent's expected utility, U, and the power of incentives, Δu, is equivalent to offering the initial contract $\langle a, b\rangle$.

To get a well-behaved model, we assume that the indirect disutility ψ satisfies Inada's conditions and has a convex marginal indirect disutility.

Assumption 2. *ψ is such that* $\lim_{\rho\to 0} \psi_1(\rho, \theta) = 0, \lim_{\rho\to 1} \psi_1(\rho, \theta) = \infty$, *and* $\psi_{111}(\rho, \theta) \geq 0$.

At this stage, it is important to stress that not all the conditions stated in this model are sufficient to ensure that the problem is concave. This difficulty is inherent to moral hazard. Nevertheless, we are able to show that if the moral-hazard problem is concave, then the mixed problem is as well.

With respect to backward induction and to the timing of the game, it is relevant to structure the rest of the paper as follows: full information, moral hazard alone (i.e., asymmetric information about action, symmetric information about ability) and mixed model (i.e., asymmetric information about both action and ability).[7]

[7] Therefore, we do not study the adverse-selection alone setup, i.e., symmetric information about action.

4. Full Information

When information is complete, the principal observes the agent's ability and action. Only the participation constraint needs to be satisfied. Given (6), the problem is $\max_{(\rho,\Delta u,U)}(4)$ subject to the participation constraint:

$$U \geq U_0. \tag{7}$$

This constraint ensures that the agent is not forced to accept the contract.

We get the following lemma.

Lemma 3. *The first-best contract entails,* $\forall\theta \in [\underline{\theta}, \overline{\theta}]$*:*

- $U^{FB}(\theta) = U_0,$
- $\Delta u^{FB}(\theta) = 0,$

so $\underline{u}^{FB}(\theta) = U^{FB}(\theta) + \psi(\rho^{FB}(\theta), \theta)$*, with* $\rho^{FB}(\theta)$ *given by:*

$$\Delta x = w'(\underline{u}^{FB}(\theta))\psi_1(\rho^{FB}(\theta), \theta). \tag{8}$$

Proof. See Appendix A.1. □

The interpretation is the following. To benefit from the increase in output value, Δx in (8), the principal must implement a positive probability of success. This implies that the agent incurs the indirect disutility. According to (4) and (5), the principal must offer the expected payment, $w(\underline{u}) + \rho\big(w(\Delta u + \underline{u}) - w(\underline{u})\big)$, while satisfying the participation constraint. However, it is increasing in $\underline{u}$ and Δu. It follows that U is costly (make use of (6)) as is Δu. It is optimal to set $U = U_0$ and $\Delta u = 0$. In other words, it is optimal not to give up a rent to the agent, $U = U_0$, nor a contingent contract, $\Delta u = 0$ and the power of incentives is null in the complete information contract. Instead, the agent receives a full insurance contract since $\Delta u = 0 \Leftrightarrow \overline{u} = \underline{u}$. The first-best payment is thus $w(\underline{u})$, with $\underline{u} = U_0 + \psi$ and the first-best marginal cost to induce a positive probability of success corresponds to the marginal payment $w'(\underline{u})\psi_1$ in (8).

5. Moral Hazard Alone

In this framework, the agent's ability is still observable, but the action is not. The principal faces a moral-hazard problem. The incentive question requires the agent to be obedient. Using Assumption 2, the agent faced with an incentive contract $\langle U, \Delta u\rangle$ chooses to generate the probability of success:

$$\begin{aligned} p(\Delta u, \theta) \quad &= \quad \arg\max_{\rho} \{\underline{u} + \rho\Delta u - \psi(\rho, \theta)\} \\ &\Rightarrow \Delta u = \psi_1(p(\Delta u, \theta), \theta) > 0. \end{aligned} \tag{9}$$

Equation (9) represents the moral-hazard incentive constraint. Now, the power of incentives must be strictly positive to ensure a positive action from the agent.

Thus, (6) becomes:

$$\begin{cases} \underline{u} = U + \psi(p(\Delta u, \theta), \theta) - p(\Delta u, \theta)\Delta u, \\ \bar{u} = U + \psi(p(\Delta u, \theta), \theta) + \big(1 - p(\Delta u, \theta)\big)\Delta u. \end{cases} \tag{10}$$

Given (10), the problem is $\max_{(\Delta u, U)}(4)$ subject to (7).

We can state the following lemma.

Lemma 4. *The moral-hazard contract entails:*

- $U^{MH}(\theta) = U_0,$ (11)
- $\Delta u^{MH}(\theta)$ *such that*

$$\Delta x = \big(w(\Delta u^{MH}(\theta) + \underline{u}^{MH}(\theta)) - w(\underline{u}^{MH}(\theta))\big) \\ + p(\Delta u^{MH}(\theta), \theta)\big(1 - p(\Delta u^{MH}(\theta), \theta)\big) \times \\ \big(w'(\Delta u^{MH}(\theta) + \underline{u}^{MH}(\theta)) - w'(\underline{u}^{MH}(\theta))\big)\psi_{11}(p(\Delta u^{MH}(\theta), \theta), \theta), \tag{12}$$

with

$$\underline{u}^{MH}(\theta) = U^{MH}(\theta) + \psi(p(\Delta u^{MH}(\theta), \theta), \theta) - p(\Delta u^{MH}(\theta), \theta)\Delta u^{MH}(\theta). \tag{13}$$

Proof. See Appendix A.2. □

This lemma deserves some comments because the marginal expected payment on the right-hand side of (12) differs from complete information. It is composed of two terms. To induce a positive action despite its non-observability, the power of incentives can no longer be set to zero, i.e., $\Delta u > 0$ (see (9)). This forces the principal to offer a contingent contract or high-powered incentives to the agent. Thus, the principal must give up the bonus to the agent, or the contingent component of the contract. This is the first term on the right-hand side of (12) (see Lemma 2). It represents the high-powered incentives marginal cost (HPIMC).

In parallel, as with complete information, the agent does not get a positive rent, $U = U_0$ in (11), because it is costly for the principal. However, because the contract is contingent, the agent bears some risk. It follows that the moral-hazard expected payment, or the moral-hazard cost to induce a positive probability of success, is higher than the first-best payment. Indeed, for a given ρ, we have:

$$\begin{aligned} w(\underline{u}) + \rho\big(w(\overline{u}) - w(\underline{u})\big) &= \big(1-\rho\big)w(\underline{u}) + \rho w(\overline{u}) \\ &> w((1-\rho)\underline{u} + \rho\overline{u}), \text{ by Jensen's inequality} \\ &= w(\underline{u} + \rho\Delta u) \\ &= w(U_0 + \psi), \text{ because } U = U_0 \text{ and } \Delta u = 0 \text{ with complete information.} \end{aligned}$$

This higher payment arises because the principal has to pay a risk-premium to the agent to ensure that the participation constraint is binding despite the risk borne. The second term on the right-hand side of (12) is the risk-premium marginal cost (RPMC), given that $\Delta u = \psi_1$ from (9).

All in all, the right-hand side of (12) is the marginal expected payment, or the moral-hazard marginal cost to induce a positive probability of success. It differs from the first-best. Therefore, the efficient action is distorted to take into account the fact that a contingent contract implies a moral-hazard cost. This is the usual insurance-efficiency trade-off in moral-hazard problems with a risk-averse agent.

6. Mixed Model

In this framework, the principal observes neither the agent's ability, nor the agent's action. Following Myerson (1982) [13], there is no loss of generality in focusing on direct revelation mechanisms. Incentives need an obedient and honest agent. However, it is convenient to recast the mixed problem as an adverse-selection problem in which the moral-hazard incentives have been inserted. To do this, let $\langle U(\hat{\theta}), \Delta u(\hat{\theta})\rangle$ be the contract offered by the principal where $\hat{\theta}$ is the agent's report on his ability.

From the moral-hazard section, we know that the agent chooses to generate a probability of success such that:

$$\begin{aligned} p(\Delta u(\hat{\theta}), \theta) &= \arg\max_{\rho}\{\underline{u}(\hat{\theta}) + \rho\Delta u(\hat{\theta}) - \psi(\rho, \theta)\} \\ &\Rightarrow \Delta u(\hat{\theta}) = \psi_1(p(\Delta u(\hat{\theta}), \theta), \theta) > 0. \end{aligned} \tag{14}$$

Let us denote by $v(\hat{\theta},\theta)$ the indirect expected utility of an agent with ability θ who reports $\hat{\theta}$. We have:

$$v(\hat{\theta},\theta) = \underline{u}(\hat{\theta}) + p(\Delta u(\hat{\theta}),\theta)\Delta u(\hat{\theta}) - \psi(p(\Delta u(\hat{\theta}),\theta),\theta).$$

Hence, the incentive constraint is $\forall\hat{\theta},\theta \in \Theta$:

$$U(\theta) = v(\theta,\theta) \geq v(\hat{\theta},\theta). \tag{15}$$

That is, the agent is better off reporting the truth about his ability.

The participation constraint is $\forall\theta \in \Theta$:

$$U(\theta) \geq U_0. \tag{16}$$

Moreover, (10) becomes:

$$\begin{cases} \underline{u}(\theta) = U(\theta) + \psi(p(\Delta u(\theta),\theta),\theta) - p(\Delta u(\theta),\theta)\Delta u(\theta), \\ \bar{u}(\theta) = U(\theta) + \psi(p(\Delta u(\theta),\theta),\theta) + \big(1 - p(\Delta u(\theta),\theta)\big)\Delta u(\theta). \end{cases} \tag{17}$$

Given (17) and using (4), the principal's problem is:

$$\max_{\Delta u(.),U(.)} \int_{\underline{\theta}}^{\overline{\theta}} \left\{\underline{x} + p(\Delta u(\theta),\theta)\Delta x - w(\underline{u}(\theta)) - p(\Delta u(\theta),\theta)\big(w(\overline{u}(\theta)) - w(\underline{u}(\theta))\big)\right\} f(\theta)d\theta, \tag{18}$$

subject to (15) and (16).

Let us begin to resolve this problem by a reformulation. The following lemma characterizes necessary and sufficient conditions for (15).

Lemma 5. *(Ollier and Thomas (2013)) [9]. The allocation $\langle U(\theta),\Delta u(\theta)\rangle$ is incentive compatible if and only if, $\forall\theta \in \Theta$:*

$$U'(\theta) = -\psi_2(p(\Delta u(\theta),\theta),\theta), \tag{19}$$

$$\Delta u'(\theta) \geq 0. \tag{20}$$

To ensure revelation, the constraint (19) tells us that the agent's expected utility, $U(\theta)$, must follow the path $U'(\theta) = -\psi_2(p(\Delta u(\theta),\theta),\theta)$. Thus, using Assumption 1, it is increasing with ability. It specifies how the informational rent must change with θ to compensate for the incentives to misreport θ if the moral-hazard contract were offered to the agent. Moreover, following (20), the principal must ensure that the power of incentives, $\Delta u(\theta)$, increases with the type. This is the implementability condition.

Next, we already know that U is costly for the principal. Since it is increasing with the agent's ability, it is optimal for the principal not to give up an informational rent to the least able agent. That is:

$$U(\underline{\theta}) = U_0. \tag{21}$$

Finally, the principal's problem is $\max_{(\Delta u(.),U(.))}$(18) s.t. (19), (20) and (21). This is an optimal control problem, where $U(\theta)$ and $\Delta u(\theta)$ are state variables.

For ease of analysis, we begin by assuming a fully separating contract, i.e., $\Delta u'(\theta) > 0, \forall\theta \in \Theta$. This allows us to identify the first-round effects of adding adverse selection to moral hazard. We study the possibility of pooling in a subsequent step.

6.1. Fully-Separating Contract

We can present our first result.

Proposition 1. *Assume a fully-separating contract. The mixed contract entails:*

- $$U^*(\theta) = U_0 - \int_{\underline{\theta}}^{\theta} \psi_2(p(\Delta u^*(\tau),\tau),\tau)d\tau, \tag{22}$$
- $\Delta u^*(\theta)$ *such that*
$$\begin{aligned}\Delta x = & \big(w(\Delta u^*(\theta) + \underline{u}^*(\theta)) - w(\underline{u}^*(\theta))\big) \\ & + p(\Delta u^*(\theta),\theta)\big(1 - p(\Delta u^*(\theta),\theta)\big) \times \\ & \quad \big(w'(\Delta u^*(\theta) + \underline{u}^*(\theta)) - w'(\underline{u}^*(\theta))\big)\psi_{11}(p(\Delta u^*(\theta),\theta),\theta) \\ & \quad - \frac{\int_\theta^{\overline{\theta}} \{w'(\underline{u}^*(\tau)) + p(\Delta u^*(\tau),\tau)\big(w'(\Delta u^*(\tau) + \underline{u}^*(\tau)) - w'(\underline{u}^*(\tau))\big)\} f(\tau)d\tau}{f(\theta)} \times \\ & \quad\quad \psi_{12}(p(\Delta u^*(\theta),\theta),\theta),\end{aligned} \tag{23}$$
with
$$\underline{u}^*(\theta) = U^*(\theta) + \psi(p(\Delta u^*(\theta),\theta),\theta) - p(\Delta u^*(\theta),\theta)\Delta u^*(\theta). \tag{24}$$

Proof. See Appendix A.3. □

Several comments can be made.

Mixed-model marginal cost. When the principal does not observe the agent's type, she needs to elicit truthful information. To do this, she must give up an informational rent to the agent, $U > U_0$, except at $\underline{\theta}$ (see (22)). Such a rent implies two modifications compared to moral hazard alone. First, this rent corresponds to the usual adverse-selection cost and is added to the moral-hazard cost to form the mixed-model cost. Second, it influences the moral-hazard cost because U is higher than U_0. Let us examine the consequences of the informational rent for the mixed-model marginal cost, i.e., the right-hand side of (23).

The first change is the presence of the last term in (23), which reflects the informational-rent marginal cost, or the adverse-selection marginal cost, and is usual in hidden information problems. It is composed of two factors. Because U is costly and must satisfy (19), i.e., $U' = -\psi_2$, it is optimal to moderate its slope, through the probability of success, to reduce the informational rent. This corresponds to ψ_{12}. The second factor is the shadow cost of U weighted by $\frac{1}{f}$. Indeed, if an agent with ability θ is offered a contract, the principal must leave an informational rent to all agents with higher ability, and so increase their expected payment, $w(\underline{u}) + \rho\big(w(\overline{u}) - w(\underline{u})\big)$ (make use of (17)), otherwise they would mimic θ. Thus, the shadow cost is $-\int_\theta^{\overline{\theta}}\{w'(\underline{u}) + p(w'(\Delta u + \underline{u}) - w'(\underline{u}))\}fd\tau$.[8]

The second modification arises because U is higher than U_0 (see (22)). Other things being equal, this increases $\underline{u}$ in (24), which in turn modifies the moral-hazard marginal cost (the first two terms in (23)). To evaluate this influence, let us proceed in accordance with the bullet points below (ignoring arguments for simplicity).

- Let $\Delta\tilde{u}$ be the solution of (23) when U is set equal to U_0 in the moral-hazard marginal cost. Furthermore, let $\tilde{p}$ and $\tilde{U}$ be the resulting probability of success and informational rent.

[8] Since the agent is not risk-neutral, the shadow cost does not reduce to $-\int_\theta^{\overline{\theta}} fd\tau = -(1-F)$, as with risk-neutrality in the canonical adverse-selection model.

- Since $\Delta\tilde{u} > 0$, the moral-hazard marginal cost can be rewritten as follows, for any $U \in [U_0, \tilde{U}]$,

$$\begin{aligned}&\left(w(\Delta\tilde{u}+\underline{u}) - w(\underline{u})\right) + \tilde{p}(1-\tilde{p})\left(w'(\Delta\tilde{u}+\underline{u}) - w'(\underline{u})\right)\psi_{11} \\ &= \int_0^{\Delta\tilde{u}} \{w'(U+\epsilon+\psi-\tilde{p}\Delta\tilde{u}) + \tilde{p}(1-\tilde{p})w''(U+\epsilon+\psi-\tilde{p}\Delta\tilde{u})\psi_{11}\}d\epsilon.\end{aligned} \tag{25}$$

Notice that $\Delta\tilde{u}$ influences the HPIMC through w' and the RPMC through w''.

- The difference between a moral-hazard marginal cost with and without an informational rent is, using (25) and the mean-value theorem,

$$(\tilde{U} - U_0)\int_0^{\Delta\tilde{u}} \{w''(\gamma+\epsilon+\psi-\tilde{p}\Delta u) + \tilde{p}(1-\tilde{p})w'''(\gamma+\epsilon+\psi-\tilde{p}\Delta\tilde{u})\psi_{11}\}d\epsilon, \tag{26}$$

with $\gamma \in [U_0, \tilde{U}]$. Thus, if (26) is negative (resp. positive), the moral-hazard marginal cost with an informational rent is lower (resp. higher) than without.

- Let us study this equation more closely. The informational rent influences the HPIMC through w'' and the RPMC through w'''. Consider first the HPIMC. The effect depends on w'', or equivalently on $-u''$ using Lemma 1 since $u' > 0$. This is positive. Indeed, due to risk aversion, the bonus incentivizing the agent to bear a risk is even more costly when the agent obtains a higher utility in the event of failure. However, using (13) and (24), the informational rent increases such utility compared with moral hazard alone. Thus, the informational rent implies that the HPIMC is increased.

  Then, consider the RPMC. Since $\psi_{11} > 0$ from Assumption 1, the effect of the informational rent depends on the sign of w''', or equivalently on $3r_A - r_P$ using Lemma 1. Therefore, the agent's risk aversion and prudence have a key role. In fact, a rent is equivalent to transferring the risk borne by the agent toward higher utility levels than the reservation utility. This constrains the principal to pay a risk premium that guarantees a higher utility in the mixed model than in the moral-hazard model alone. At the margin, this is costly for the principal because of risk aversion. The term $3r_A$ reflects this phenomenon. The role of prudence is more complex. According to decision theory, the transfer of the risk toward higher levels of utility is not welcomed by an imprudent agent, but the contrary is true for a prudent one. Therefore, in the former case, i.e., $-r_P > 0$, imprudence makes the transfer of risk costly, as does risk aversion. The RPMC increases in the presence of the informational rent. In the latter case, i.e., $-r_P < 0$, risk aversion and prudence work in opposite directions. Thus, prudence reduces the RPMC contrarily to risk aversion. The ultimate effect depends on the level of prudence. Using Definition 1, when prudence is low, risk aversion dominates prudence and the informational rent still increases the RPMC because $3r_A - r_P$ remains positive. However, when prudence is high, the informational rent decreases the RPMC since $3r_A - r_P < 0$.

- The overall effect of the informational rent on the moral-hazard marginal cost depends on the sign of the integrand in (26) or equivalently using Lemma 1, of (since $u' > 0$):

$$-(u''u'^2 + \tilde{p}(1-\tilde{p})(u'''u' - 3u''^2)\psi_{11})$$

or, by factoring $-u''u'$ and making use of Assumption 1, Definition 1 and $u''u' < 0$,

$$\frac{u'}{\tilde{p}(1-\tilde{p})\psi_{11}} + 3r_A - r_P. \tag{27}$$

- This equation allows us to introduce the following definition.

Definition 2. *Consider the contract $\langle\tilde{U}, \Delta\tilde{u}\rangle$. Since $\frac{u'}{\tilde{p}(1-\tilde{p})\psi_{11}} > 0$ at an interior solution, an agent is said sufficiently highly prudent if (27), and thus, (26) is negative.*

Finally, the effect of the informational rent on the moral-hazard marginal cost in the mixed model can be summarized in the three following points:

a. if the agent is imprudent, (26) is positive. Risk aversion and prudence both contribute to increasing the moral-hazard marginal cost,
b. if the agent is weakly or highly but not sufficiently prudent, (26) is still positive. However, prudence softens the increase in moral-hazard marginal cost due to risk aversion,
c. if the agent is sufficiently highly prudent, (26) is negative. Therefore, a reduction in the moral-hazard marginal cost occurs because prudence offsets risk aversion.

Altogether, the right-hand side of (23) is the mixed-model marginal cost, i.e., the sum of the moral-hazard marginal cost and the adverse-selection marginal cost. When it equals the marginal benefit, the optimal probability is obtained.

Distortions with respect to moral hazard. These two modifications imply distortions in the probability of success, compared to its moral-hazard level. They are stated in the following proposition.

Proposition 2. *Consider a sufficiently highly prudent agent. The mixed-model probability of success $p(\Delta u^*(\theta), \theta)$ is higher than the moral-hazard probability $p(\Delta u^{MH}(\theta), \theta)$ if the agent is sufficiently able, that is, an agent for whom the decrease in the moral-hazard marginal cost is higher than the adverse-selection marginal cost.*
Otherwise, the mixed probability is lower.

To see this, notice that (23) is equivalent to:

$$\begin{aligned}\Delta x + & \frac{\int_{\theta}^{\overline{\theta}} \{w'(\underline{u}) + p\big(w'(\overline{u}) - w'(\underline{u})\big)\} f d\tau}{f(\theta)} \psi_{12}(p,\theta) \\ & = \big(w(\overline{u}) - w(\underline{u})\big) + p(1-p)\big(w'(\overline{u}) - w'(\underline{u})\big)\psi_{11}(p,\theta). \end{aligned} \tag{28}$$

We know that $\Delta x + \frac{\int_{\theta}^{\overline{\theta}} \{w'(\underline{u}) + p(w'(\Delta u + \underline{u}) - w'(\underline{u}))\} f d\tau}{f} \psi_{12} \leq \Delta x$, with equality holding at $\overline{\theta}$. Therefore, comparing (12) and (28), the first modification forces a reduction in p because the moral-hazard marginal cost is increasing in p.[9] This reduction reflects the usual rent extraction-efficiency trade-off in pure adverse-selection problems: this tends to distort the probability downward from its reference level (here its moral-hazard level) in order to reduce the informational rent left to more able agents. Observe that this cost is null at $\overline{\theta}$ or for abilities close to this type.

Is this decrease in the probability of success sufficient to be optimal? The preceding paragraph shows that it is not, because the moral-hazard marginal cost is modified by the informational rent, i.e., $U > U_0$, except at $\underline{\theta}$. As shown above, this can increase (Cases a. and b.) or reduce (Case c.) the moral-hazard marginal cost. In Cases a. and b., adverse selection contributes to a second reduction in p because the transfer of risk toward higher levels of utility involved makes the insurance-efficiency trade-off worse. In Case c., adverse selection leads to an opposite effect and tends to increase p. In this case, the informational rent combined with a sufficiently high level of prudence softens the trade-off due to moral hazard.

These distortions reflect the overall trade-off between rent extraction, insurance, and efficiency. Finally, the proposition says that if the informational rent leads to a reduction in the moral-hazard marginal cost that is higher than the adverse-selection marginal cost, then the likelihood of success is higher under both adverse selection and moral hazard than under moral hazard alone. This occurs if the agent is sufficiently highly prudent and sufficiently able. In all other cases, the probability of success is lower.

---

9 By concavity of the moral-hazard problem, the moral-hazard marginal cost increases in p. See Equation (A10) in Appendix A.2.

Moral-hazard cost versus adverse-selection cost. From the preceding discussion, it follows that the nature of distortions is not equivalent among $[\underline{\theta}, \overline{\theta}]$. The following proposition provides details about this phenomenon.

Proposition 3. *Moving from the lowest to the highest agent's ability, the mixed contract replaces distortions due to adverse-selection cost by distortions due to moral-hazard cost. Thus, there is also a distortion at the top.*

Indeed, when θ is close to $\underline{\theta}$, we have $U \simeq U_0$ and $-\frac{\int_{\theta}^{\overline{\theta}}\{w'(\underline{u})+p(w'(\Delta u+\underline{u})-w'(\underline{u}))\}fd\tau}{f}\psi_{12} \gg 0$. Therefore, the distortion associated with low types is due rather to the rent extraction-efficiency trade-off. The principal is more concerned about the adverse-selection cost. The probability of success is distorted to limit the informational rent. The reverse is true for θ close to $\overline{\theta}$ since $U \gg U_0$ and $-\frac{\int_{\theta}^{\overline{\theta}}\{w'(\underline{u})+p(w'(\Delta u+\underline{u})-w'(\underline{u}))\}fd\tau}{f}\psi_{12} \simeq 0$. The distortion associated with able types occurs rather from the insurance-efficiency trade-off, because the principal is more interested in the moral-hazard cost. The distortion arises from the desire to limit the cost to induce a positive probability of success, i.e., the high-powered incentives and the risk-premium. It follows that in presence of adverse selection, each moral-hazard probability of success is distorted. In particular, even if the marginal cost of the informational rent is null for the highest type, because $\frac{\int_{\theta}^{\overline{\theta}}\{w'(\underline{u})+p(w'(\Delta u+\underline{u})-w'(\underline{u}))\}fd\tau}{f} = 0$ at $\theta = \overline{\theta}$, there is a distortion at the top.

6.2. Partially Separating Contract

As just seen, the aim of the distortions varies with the agent's type. This raises the question whether their effects and the agent's ability covary in the same direction. In other words, this raises the question whether the assumption of full separation in Proposition 1 is relevant. Let $\eta(\theta)$ represent the shadow cost of U, $\eta(\theta) = -\int_{\theta}^{\overline{\theta}}\{w'(\underline{u}) + p(w'(\Delta u + \underline{u}) - w'(\underline{u}))\}fd\tau$. The following proposition answers this question.

Proposition 4. *The optimal mixed contract is separating if,* $\forall \theta \in [\underline{\theta}, \overline{\theta}]$:

$$\begin{aligned} &p_{12}(\Delta u^*(\theta),\theta)\big(\Delta x - \big(w(\overline{u}^*(\theta)) - w(\underline{u}^*(\theta))\big)\big) \\ &\quad - p_2(\Delta u^*(\theta),\theta)\big(1-2p(\Delta u^*(\theta),\theta)\big)\big(w'(\overline{u}^*(\theta)) - w'(\underline{u}^*(\theta))\big) \\ &\quad + p_2(\Delta u^*(\theta),\theta)\left(\frac{\eta^*(\theta)}{f}\right)' + p_{22}(\Delta u^*(\theta),\theta)\frac{\eta^*(\theta)}{f(\theta)} > 0. \end{aligned} \tag{29}$$

Proof. See Appendix A.4. □

Two comments can be made. First, this proposition highlights that pooling in this mixed framework can have two sources. The first source arises due to non-responsiveness. That is, $\Delta u^{MH}(\theta)$ does not satisfy the incentive constraint (20). In this case, the principal would like to implement a power of incentives decreasing in θ for reasons of moral hazard. Thus, a conflict appears with the implementability condition. The first two terms in (29) reflect this phenomenon. However, because $\underline{u}$ depends on $U > U_0$ when there is adverse selection, the conflict between the principal's preference and the monotonicity condition (20) is somewhat modified compared to moral hazard alone. The second source is due to the possible lack of monotonicity of the marginal cost of the informational rent. In such a situation, this cost is not ranked exactly as the agent's ability and the principal would like to implement a decreasing power of incentives over some interval. Again, this conflicts with the implementability condition. This concerns the final two terms in (29). This effect is well-known in pure adverse-selection models. However, because the agent is risk-averse, the usual monotone hazard rate property, i.e., $\frac{1-F}{f}$ non-increasing in θ, is no longer sufficient to avoid non-monotonicity.

Second, examining (29), we observe that since $p_2 = -\frac{\psi_{12}}{\psi_{11}} > 0$ and $w'' > 0$ (see Lemma 1), a probability less than $\frac{1}{2}$ contributes to the non-responsiveness of the model. The reverse is true if $p \geq \frac{1}{2}$. Moreover, since $\Delta x - \big(w(\overline{u}) - w(\underline{u})\big) > 0$, $\eta' > 0$ and $\eta < 0$, (29) is more easily satisfied when:

$$p_{12} \geq 0 \text{ and } p_{22} \leq 0. \tag{30}$$

These conditions structure the function ψ even more than Assumptions 1 and 2 do. Indeed, we get:

$$\begin{aligned} p_{12}(\Delta u,\theta) &= \frac{\psi_{111}(p(\Delta u,\theta),\theta)\psi_{12}(p(\Delta u,\theta),\theta) - \psi_{112}(p(\Delta u,\theta),\theta)\psi_{11}(p(\Delta u,\theta),\theta)}{\psi_{11}^3(p(\Delta u,\theta),\theta)}, \\ p_{22}(\Delta u,\theta) &= \frac{\psi_{112}(p(\Delta u,\theta),\theta)\psi_{12}(p(\Delta u,\theta),\theta) - \psi_{212}(p(\Delta u,\theta),\theta)\psi_{11}(p(\Delta u,\theta),\theta)}{\psi_{11}^2(p(\Delta u,\theta),\theta)} \\ &\quad - \frac{\psi_{12}(p(\Delta u,\theta),\theta)}{\psi_{11}(p(\Delta u,\theta),\theta)} \frac{\psi_{111}(p(\Delta u,\theta),\theta)\psi_{12}(p(\Delta u,\theta),\theta) - \psi_{112}(p(\Delta u,\theta),\theta)\psi_{11}(p(\Delta u,\theta),\theta)}{\psi_{11}^2(p(\Delta u,\theta),\theta)}. \end{aligned}$$

Therefore, a separating contract is subject to a combination of many second and third partial derivatives of the indirect disutility ψ. Thus, except maybe for very simple functions ψ, one can reasonably have doubts about the existence of a contract that strictly satisfies (20) for all types. In this case, the monotonicity constraint is binding on some intervals.

We state the shape of the optimal mixed contract in the following proposition.

Proposition 5. *Consider a single interior interval $[\theta_0, \theta_1]$ where there is pooling. The mixed contract entails:*

- $U^{**}(\theta) = U_0 - \int_{\underline{\theta}}^{\theta} \psi_2(p(\Delta u^{**}(\tau),\tau),\tau)d\tau,$
- $\Delta u^{**}(\theta)$ *equal to*
  - $\Delta u^*(\theta)$ *if* $\theta \in [\underline{\theta}, \theta_0] \cup [\theta_1, \overline{\theta}]$,
  - Δu^k *if* $\theta \in [\theta_0, \theta_1]$, *with*

$$\begin{aligned} \Delta x - \int_{\theta_0}^{\theta_1} \Big\{ &\big(w(\Delta u^k + \underline{u}^{**}(\theta)) - w(\underline{u}^{**}(\theta))\big) \\ &- p(\Delta u^k,\theta)\big(1 - p(\Delta u^k,\theta)\big) \times \\ &\quad \big(w'(\Delta u^k + \underline{u}^{**}(\theta)) - w'(\underline{u}^k(\theta))\big)\psi_{11}(p(\Delta u^k,\theta),\theta) \\ &+ \frac{\int_{\theta}^{\overline{\theta}} \{w'(\underline{u}^{**}(\tau)) + p(\Delta u^k,\theta)\big(w'(\Delta u^k + \underline{u}^{**}(\tau)) - w'(\underline{u}^{**}(\tau))\big)\} f(\tau)d\tau}{f(\theta)} \times \\ &\quad \psi_{12}(p(\Delta u^k,\theta),\theta) \Big\} f(\theta)d\theta = 0 \end{aligned}$$

and $\quad \underline{u}^{**}(\theta) = U^{**}(\theta) + \psi(p(\Delta u^{**}(\theta),\theta),\theta) - p(\Delta u^{**}(\theta),\theta)\Delta u^{**}(\theta).$

This proposition shows that when pooling arises, the optimal contract consists in verifying (23) on average, and no longer pointwise. This result is well-known in pure adverse-selection models.

In the light of Propositions 4 and 5, we observe that pooling in the mixed contract has no different causes and consequences from those already known in pure adverse-selection models. However, pooling is most likely to emerge since the conditions to avoid pooling are more difficult to satisfy because of the different kinds of incentives the principal has to manage and the overall trade-off between rent extraction, insurance, and efficiency involved.

7. An Example

It is important to notice that the optimal power of incentives Δu^* defined in Proposition 1 is the solution of a non-linear integral state equation. Thus, the aim of this example is not to find an explicit solution but to give a simple setting in order to show: (1) the existence of the distortion at the top; (2) the existence of a sufficiently highly prudent agent and of an increase in the probability of success in the mixed model; and (3) whether the conditions ensuring a separating contract can arise.

To do this, let $x = u(y) = (\alpha y)^{\frac{1}{\alpha}}$, with $1 < \alpha \leq 2$. We get:

$$u' = \alpha^{\frac{1-\alpha}{\alpha}} y^{\frac{1-\alpha}{\alpha}}; u'' = \frac{1-\alpha}{\alpha}\alpha^{\frac{1-\alpha}{\alpha}} y^{\frac{1-2\alpha}{\alpha}}; u''' = \frac{1-2\alpha}{\alpha}\frac{1-\alpha}{\alpha}\alpha^{\frac{1-\alpha}{\alpha}} y^{\frac{1-3\alpha}{\alpha}},$$

so

$$r_A = -\frac{1-\alpha}{\alpha}y^{-1}, r_P = -\frac{1-2\alpha}{\alpha}y^{-1},$$

and

$$r_P - 3r_A = \frac{2-\alpha}{\alpha}y^{-1}. \tag{31}$$

Moreover, we obtain:

$$y = w(x) = \frac{1}{\alpha}x^{\alpha}; w'(x) = x^{\alpha-1}; w''(x) = (\alpha-1)x^{\alpha-2}; w'''(x) = (\alpha-1)(\alpha-2)x^{\alpha-3}.$$

Then, let $\psi(\rho,\theta) = \frac{\rho^2}{\theta}$. It follows immediately that $\psi_1 = \frac{2\rho}{\theta}$, $\psi_2 = -\frac{\rho^2}{\theta^2}$, $\psi_{11} = \frac{2}{\theta}$ and $\psi_{12} = -\frac{2\rho}{\theta^2}$.[10] Finally, let $p(\theta) = p(\Delta u(\theta),\theta)$ to simplify.

7.1. Distortion at the Top

In this subsection, we let $\alpha = 2$. Therefore, we obtain $y = w(x) = \frac{x^2}{2}$, $w'(x) = x$ and $w''(x) = 1$.

Full information. Since $U = U_0$, $\Delta u = 0$ and $\underline{u} = \psi + U_0$, (8) becomes:

$$\Delta x = \left(U_0 + \frac{\rho^{FB}(\theta)^2}{\theta}\right)\frac{2\rho^{FB}(\theta)}{\theta}. \tag{32}$$

Moral hazard alone. We have:

$$\begin{aligned} w(\Delta u + \underline{u}) - w(\underline{u}) &= (U_0+\psi)\psi_1 + \frac{1-2\rho}{2}\psi_1^2 \\ &= \left(U_0 + \frac{\rho^2}{\theta}\right)\frac{2\rho}{\theta} + \frac{1-2\rho}{2}\left(\frac{2\rho}{\theta}\right)^2, \\ w'(\Delta u + \underline{u}) - w'(\underline{u}) &= \psi_1 = \frac{2\rho}{\theta}. \end{aligned}$$

Therefore, (12) becomes:

$$\begin{aligned} \Delta x = &\left(U_0 + \frac{p^{MH}(\theta)^2}{\theta}\right)\frac{2p^{MH}(\theta)}{\theta} \\ &+ \frac{1-2p^{MH}(\theta)}{2}\left(\frac{2p^{MH}(\theta)}{\theta}\right)^2 + p^{MH}(\theta)(1-p^{MH}(\theta))\left(\frac{2p^{MH}(\theta)}{\theta}\right)\left(\frac{2}{\theta}\right). \end{aligned} \tag{33}$$

[10] Using this utility function requires a positive payment in the event of failure. Since $\rho\psi_1 - \psi = \frac{\rho^2}{\theta} > 0$, $\Delta u = \psi_1$ from (9), and $\underline{u} = U + \psi(\rho,\theta) - \rho\Delta u$ from (6), we assume that $U_0 > \frac{\rho^2}{\theta}$ to ensure that $\underline{u}$, and so a, are positive.

Comparing (32) and (33), the distortion due to moral hazard is:

$$\frac{1-2p}{2}\left(\frac{2p}{\theta}\right)^2 + p(1-p)\left(\frac{2p}{\theta}\right)\left(\frac{2}{\theta}\right).$$

Mixed model. Since $-\int_{\underline{\theta}}^{\theta}\psi_2 d\tau = \int_{\underline{\theta}}^{\theta}\frac{\rho^2}{\tau^2}d\tau$, we have:

$$\begin{aligned}
w(\Delta u+\underline{u})-w(\underline{u}) &= (U+\psi)\psi_1 + \frac{1-2\rho}{2}\psi_1^2\\
&= \left(U_0+\int_{\underline{\theta}}^{\theta}\frac{\rho^2}{\tau^2}d\tau+\frac{\rho^2}{\theta}\right)\frac{2\rho}{\theta}+\frac{1-2\rho}{2}\left(\frac{2\rho}{\theta}\right)^2,\\
w'(\Delta u+\underline{u})-w'(\underline{u}) &= \psi_1 = \frac{2\rho}{\theta},\\
w'(\underline{u})+\rho(w'(\Delta u+\underline{u})-w'(\underline{u})) &= U+\psi\\
&= U_0+\int_{\underline{\theta}}^{\theta}\frac{\rho^2}{\tau^2}d\tau+\frac{\rho^2}{\theta}.
\end{aligned}$$

Therefore, using Proposition 1, the optimal probability of success is such that:

$$\begin{aligned}
\Delta x = &\left(U_0+\frac{p^*(\theta)^2}{\theta}\right)\frac{2p^*(\theta)}{\theta}\\
&+\frac{1-2p^*(\theta)}{2}\left(\frac{2p^*(\theta)}{\theta}\right)^2+p^*(\theta)(1-p^*(\theta))\left(\frac{2p^*(\theta)}{\theta}\right)\left(\frac{2}{\theta}\right)\\
&+\int_{\underline{\theta}}^{\theta}\frac{p^*(\tau)^2}{\tau^2}d\tau\frac{2p^*(\theta)}{\theta}\\
&+\frac{2p^*(\theta)}{\theta}\frac{\int_{\theta}^{\overline{\theta}}\left(U_0+\int_{\underline{\theta}}^{\epsilon}\frac{p^*(\epsilon)^2}{\epsilon^2}d\epsilon+\frac{p^*(\theta)^2}{\theta}\right)f(\tau)d\tau}{f(\theta)}.
\end{aligned} \tag{34}$$

If we compare (34) to (33), the last two lines of the preceding equation reflect the distortions due to adverse selection. The penultimate term is the increase in the moral-hazard marginal cost because the agent is weakly prudent (i.e., $r_P - 3r_A = 0$ when $\alpha = 2$ from (31)). The last term is the marginal cost of the informational rent. Therefore, the mixed-model probability of success is distorted twice downward from its moral-hazard level.

At $\theta = \overline{\theta}$, the last term is null but the penultimate is not. Therefore, the term $\int_{\underline{\theta}}^{\theta}\frac{p^*(\tau)^2}{\tau^2}d\tau\frac{2p^*(\theta)}{\theta}$ leads to the distortion at the top.

7.2. Sufficiently Highly Prudence

Faced with the difficulty to solve a non-linear integral equation, we adopt the following simplification. We still consider a continuous probability of success, but only two types instead of a continuum. Let $\theta \in \{\underline{\theta}, \overline{\theta}\}$ with $\underline{\theta} < \overline{\theta}$, $\underline{f} = \Pr(\theta = \underline{\theta})$ and $\overline{f} = \Pr(\theta = \overline{\theta})$. The $\overline{\theta}$-incentive constraint is:

$$\begin{aligned}
U(\overline{\theta}) = v(\overline{\theta},\overline{\theta}) &\geq \underline{u}(\underline{\theta}) + p(\Delta u(\underline{\theta}),\overline{\theta})\Delta u(\underline{\theta}) - \psi(p(\Delta u(\underline{\theta}),\overline{\theta}),\overline{\theta})\\
&= \underline{u}(\underline{\theta}) + p(\Delta u(\underline{\theta}),\underline{\theta})\Delta u(\underline{\theta}) - \psi(p(\Delta u(\underline{\theta}),\underline{\theta}),\underline{\theta})\\
&\quad + p(\Delta u(\underline{\theta}),\overline{\theta})\Delta u(\underline{\theta}) - p(\Delta u(\underline{\theta}),\underline{\theta})\Delta u(\underline{\theta})\\
&\quad + \psi(p(\Delta u(\underline{\theta}),\underline{\theta}),\underline{\theta}) - \psi(p(\Delta u(\underline{\theta}),\overline{\theta}),\overline{\theta})\\
&= U(\underline{\theta}) + p(\Delta u(\underline{\theta}),\overline{\theta})\Delta u(\underline{\theta}) - p(\Delta u(\underline{\theta}),\underline{\theta})\Delta u(\underline{\theta})\\
&\quad + \psi(p(\Delta u(\underline{\theta}),\underline{\theta}),\underline{\theta}) - \psi(p(\Delta u(\underline{\theta}),\overline{\theta}),\overline{\theta}).
\end{aligned} \tag{35}$$

Thus, the $\overline{\theta}$-informational rent is:

$$\begin{aligned} U(\overline{\theta}) - U_0 = U(\overline{\theta}) - U(\underline{\theta}) =& p(\Delta u(\underline{\theta}),\overline{\theta})\Delta u(\underline{\theta}) - p(\Delta u(\underline{\theta}),\underline{\theta})\Delta u(\underline{\theta}) \\ &+ \psi(p(\Delta u(\underline{\theta}),\underline{\theta}),\underline{\theta}) - \psi(p(\Delta u(\underline{\theta}),\overline{\theta}),\overline{\theta}). \end{aligned} \tag{36}$$

Let $\underline{p} = p(\underline{\theta})$ and $\overline{p} = p(\overline{\theta})$. Hence, according to our specification:

- the informational rent becomes[11] $\frac{\underline{p}^2}{\underline{\theta}^2}\left(\overline{\theta} - \underline{\theta}\right)$;
- by Lemma 4, the probabilities $(\underline{p}^{MH}, \overline{p}^{MH})$ are given by:

$$\begin{aligned} \Delta x =& \frac{1}{\alpha}\left(\left(U_0 + \frac{\underline{p}^{MH^2}}{\underline{\theta}} + (1-\underline{p}^{MH})\frac{2\underline{p}^{MH}}{\underline{\theta}}\right)^{\alpha} - \left(U_0 + \frac{\underline{p}^{MH^2}}{\underline{\theta}} - \underline{p}^{MH}\frac{2\underline{p}^{MH}}{\underline{\theta}}\right)^{\alpha}\right) \\ &+ \underline{p}^{MH}(1-\underline{p}^{MH})\left(\left(U_0 + \frac{\underline{p}^{MH^2}}{\underline{\theta}} + (1-\underline{p}^{MH})\frac{2\underline{p}^{MH}}{\underline{\theta}}\right)^{\alpha-1}\right. \\ &\left. - \left(U_0 + \frac{\underline{p}^{MH^2}}{\underline{\theta}} - \underline{p}^{MH}\frac{2\underline{p}^{MH}}{\underline{\theta}}\right)^{\alpha-1}\right)\frac{2}{\underline{\theta}}, \end{aligned} \tag{37}$$

and:

$$\begin{aligned} \Delta x =& \frac{1}{\alpha}\left(\left(U_0 + \frac{\overline{p}^{MH^2}}{\overline{\theta}} + (1-\overline{p}^{MH})\frac{2\overline{p}^{MH}}{\overline{\theta}}\right)^{\alpha} - \left(U_0 + \frac{\overline{p}^{MH^2}}{\overline{\theta}} - \overline{p}^{MH}\frac{2\overline{p}^{MH}}{\overline{\theta}}\right)^{\alpha}\right) \\ &+ \overline{p}^{MH}(1-\overline{p}^{MH})\left(\left(U_0 + \frac{\overline{p}^{MH^2}}{\overline{\theta}} + (1-\overline{p}^{MH})\frac{2\overline{p}^{MH}}{\overline{\theta}}\right)^{\alpha-1}\right. \\ &\left. - \left(U_0 + \frac{\overline{p}^{MH^2}}{\overline{\theta}} - \overline{p}^{MH}\frac{2\overline{p}^{MH}}{\overline{\theta}}\right)^{\alpha-1}\right)\frac{2}{\overline{\theta}}; \end{aligned} \tag{38}$$

- by Proposition 1, the probabilities $\underline{p}^*$ and $\overline{p}^*$ are given by:

$$\begin{aligned} \Delta x =& \frac{1}{\alpha}\left(\left(U_0 + \frac{\underline{p}^{*2}}{\underline{\theta}} + (1-\underline{p}^*)\frac{2\underline{p}^*}{\underline{\theta}}\right)^{\alpha} - \left(U_0 + \frac{\underline{p}^{*2}}{\underline{\theta}} - \underline{p}^*\frac{2\underline{p}^*}{\underline{\theta}}\right)^{\alpha}\right) \\ &+ \underline{p}^*(1-\underline{p}^*)\left(\left(U_0 + \frac{\underline{p}^{*2}}{\underline{\theta}} + (1-\underline{p}^*)\frac{2\underline{p}^*}{\underline{\theta}}\right)^{\alpha-1} - \left(U_0 + \frac{\underline{p}^{*2}}{\underline{\theta}} - \underline{p}^*\frac{2\underline{p}^*}{\underline{\theta}}\right)^{\alpha-1}\right)\frac{2}{\underline{\theta}} \\ &+ \frac{\overline{f}}{\underline{f}}\left[\left(U_0 + \frac{\underline{p}^{*2}}{\underline{\theta}^2}\left(\overline{\theta}-\underline{\theta}\right) + \frac{\overline{p}^{*2}}{\overline{\theta}} - \overline{p}^*\frac{2\overline{p}^*}{\overline{\theta}}\right)^{\alpha-1}\right. \\ &\quad + \overline{p}^*\left(\left(U_0 + \frac{\underline{p}^{*2}}{\underline{\theta}^2}\left(\overline{\theta}-\underline{\theta}\right) + \frac{\overline{p}^{*2}}{\overline{\theta}} + (1-\overline{p}^*)\frac{2\overline{p}^*}{\overline{\theta}}\right)^{\alpha-1}\right. \\ &\quad \left.\left. - \left(U_0 + \frac{\underline{p}^{*2}}{\underline{\theta}^2}\left(\overline{\theta}-\underline{\theta}\right) + \frac{\overline{p}^{*2}}{\overline{\theta}} - \overline{p}^*\frac{2\overline{p}^*}{\overline{\theta}}\right)^{\alpha-1}\right)\right]\frac{2\underline{p}^*}{\underline{\theta}^2}\left(\overline{\theta}-\underline{\theta}\right), \end{aligned} \tag{39}$$

and:

---

[11] From (14), we get $p(\Delta u(\underline{\theta}),\theta) = \frac{\Delta u(\underline{\theta})\theta}{2}, \forall\theta \in \{\underline{\theta},\overline{\theta}\}$. The informational rent in (36) becomes $\left(\frac{\Delta u(\underline{\theta})\overline{\theta}}{2} - \frac{\Delta u(\underline{\theta})\underline{\theta}}{2}\right)\Delta u(\underline{\theta}) + \left(\frac{\left(\frac{\Delta u(\underline{\theta})\underline{\theta}}{2}\right)^2}{\underline{\theta}} - \frac{\left(\frac{\Delta u(\underline{\theta})\overline{\theta}}{2}\right)^2}{\overline{\theta}}\right) = \frac{\Delta u(\underline{\theta})^2}{4}(\overline{\theta}-\underline{\theta})$. Using $\Delta u(\underline{\theta}) = \frac{2\underline{p}}{\underline{\theta}}$ completes the proof.

$$
\begin{aligned}
\Delta x = &\frac{1}{\alpha}\left(\left(U_0+\frac{\underline{p}^{*2}}{\underline{\theta}^2}\left(\overline{\theta}-\underline{\theta}\right)+\frac{\overline{p}^{*2}}{\overline{\theta}}+\left(1-\overline{p}^{*}\right)\frac{2\overline{p}^{*}}{\overline{\theta}}\right)^{\alpha}\right.\\
&\qquad\left.-\left(U_0+\frac{\underline{p}^{*2}}{\underline{\theta}^2}\left(\overline{\theta}-\underline{\theta}\right)+\frac{\overline{p}^{*2}}{\overline{\theta}}-\overline{p}^{*}\frac{2\overline{p}^{*}}{\overline{\theta}}\right)^{\alpha}\right)\\
&+\overline{p}^{*}(1-\overline{p}^{*})\left(\left(U_0+\frac{\underline{p}^{*2}}{\underline{\theta}^2}\left(\overline{\theta}-\underline{\theta}\right)+\frac{\overline{p}^{*2}}{\overline{\theta}}+\left(1-\overline{p}^{*}\right)\frac{2\overline{p}^{*}}{\overline{\theta}}\right)^{\alpha-1}\right.\\
&\qquad\left.-\left(U_0+\frac{\underline{p}^{*2}}{\underline{\theta}^2}\left(\overline{\theta}-\underline{\theta}\right)+\frac{\overline{p}^{*2}}{\overline{\theta}}-\overline{p}^{*}\frac{2\overline{p}^{*}}{\overline{\theta}}\right)^{\alpha-1}\right)\frac{2}{\overline{\theta}}.
\end{aligned}
\tag{40}
$$

Let:

$$\Delta x = 1, \alpha = 1.01, U_0 = 0.5, (\underline{\theta}, \overline{\theta}) = (0.4, 0.5), \text{ and } (\underline{f}, \overline{f}) = (0.3, 0.7). \tag{41}$$

Equations (37) and (38) lead to:

$$(\underline{p}^{MH}, \overline{p}^{MH}) = (0.198353, 0.248085).$$

Moreover, according to the definition of $\Delta\tilde{u}$, we get $\tilde{\overline{p}} = \overline{p}^{MH} = 0.248085$. Plugging this value in (39), we get $\tilde{\underline{p}} = 0.126235$. Using these values to compute the right-hand-side of (40), it is equal to 0.999989. This is lower than 1 (i.e., Δx). Thus, at $(\tilde{\underline{p}}, \tilde{\overline{p}})$, the $\overline{\theta}$-mixed-model marginal cost is lower than the $\overline{\theta}$-moral-hazard marginal cost. By Definition 2, $\overline{\theta}$ is sufficiently highly prudent. Since the $\overline{\theta}$-adverse-selection marginal cost is null, $\overline{\theta}$ is also sufficiently able.

Thus, by Proposition 2, we must have $\overline{p}^* > \overline{p}^{MH}$. Solving the system (39),(40) with (41), we actually find [12]

$$(\underline{p}^*, \overline{p}^*) = (0.126235, 0.248088).$$

7.3. Separating Contract

According to (14), the agent chooses to generate a probability of success such that:

$$p(\Delta u(\hat{\theta}), \theta) = \frac{\Delta u(\hat{\theta})\theta}{2}. \tag{42}$$

It follows immediately that $p_{12} = \frac{1}{2}$ and $p_{22} = 0$. Thus, the conditions in (30) are satisfied. They contribute to obtaining a separating contract.

8. Conclusions

In this paper, we have studied the contract between a risk-neutral principal and a risk-averse agent who has private information about his ability and action that both improve the probability of success. In a two-output model, we have shown that if the agent is sufficiently highly prudent and able, the principal induces a higher probability of success than under moral hazard, despite the costly informational rent given up. Moreover, there is inevitably distortion at the top. Finally, the conditions for avoiding pooling are difficult to satisfy because of the different kinds of incentives to be managed and the overall trade-off between rent extraction, insurance, and efficiency involved.

Two natural extensions would be interesting. The first one would be to consider more than two outputs, even a continuum. The complexity arises from finding a tractable form of the moral-hazard marginal cost to be able to analyze the influence of the informational rent on the moral-hazard trade-off. Secondly, this two-output model can be used to investigate an insurance relationship. However, such a

12 More precisely, we have $\tilde{\underline{p}} < \underline{p}^*$ since $\tilde{\underline{p}} = 0.12623470368$ and $\underline{p}^* = 0.12623470381$.

contract must consider an outside option corresponding to the expected utility obtained by the agent when he does not purchase insurance. The difficulty is thus to measure the influence of the outside option on the contract offered by the principal.

Acknowledgments: The authors would like to acknowledge financial support by institutions through the Knowledge Unlatched initiative for the Open Access publishing. We are also very grateful to three referees and the Editor Pak Hung Au for very helpful comments. All errors remain, of course, the sole responsibility of the authors.

Author Contributions: The model was initiated by Lionel Thomas and developed and finalized by François Maréchal and Lionel Thomas

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

For the sake of simplicity in the appendix, we focus on interior solutions.

Appendix A.1. Proof of Lemma 3

In this informational setting, an interior solution means that $\Delta u \in \mathbb{R}$ and $\rho \in (0,1)$. We denote μ the Kuhn and Tucker multiplier associated to (7). The Lagrangian is:

$$\begin{aligned} L = {} & \underline{x} + \rho\Delta x - w(U + \psi - \rho\Delta u) \\ & - \rho(w(U + \psi + (1-\rho)\Delta u) - w(U + \psi - \rho\Delta u))) + \mu(U - U_0). \end{aligned} \tag{A1}$$

Necessary conditions are:

$$\begin{aligned} \frac{\partial L}{\partial \rho} = {} & \Delta x - w'(\underline{u})(\psi_1 - \Delta u) - (w(\Delta u + \underline{u}) - w(\underline{u})) \\ & - \rho(w'(\Delta u + \underline{u})(\psi_1 - \Delta u) - w'(\underline{u})(\psi_1 - \Delta u)) = 0 \end{aligned} \tag{A2}$$

$$\frac{\partial L}{\partial U} = -w'(\underline{u}) - \rho(w'(\Delta u + \underline{u}) - w'(\underline{u})) + \mu = 0 \tag{A3}$$

$$\frac{\partial L}{\partial \Delta u} = w'(\underline{u})\rho - \rho(w'(\Delta u + \underline{u})(1-\rho) + w'(\underline{u})\rho) = 0 \tag{A4}$$

$$\mu \geq 0, \mu(U - U_0) = 0. \tag{A5}$$

From (A4), we have:

$$w'(\Delta u + \underline{u}) = w'(\underline{u}) \Rightarrow \Delta u = 0.$$

Then, plugging this result into (A3), we have $\mu = w'(\underline{u}) > 0$ using Lemma 1. Therefore, from (A5), we get $U = U_0$. After simplifications in (A2), the probability $\rho^{FB}(\theta)$ is given by (8).

Appendix A.2. Proof of Lemma 4

Looking for $\rho \in (0,1)$, the Lagrangian is similar to (A1),

$$\begin{aligned} L \quad = \quad & \underline{x} + p\Delta x - w(U + \psi - p\Delta u) \\ & -p(w(U + \psi + (1-p)\Delta u) - w(U + \psi - p\Delta u)) + \mu(U - U_0). \end{aligned}$$

Necessary conditions. Given (9), the necessary conditions are:

$$\frac{\partial L}{\partial U} = -w'(\underline{u}) - p(w'(\Delta u + \underline{u}) - w'(\underline{u})) + \mu = 0 \tag{A6}$$

$$\frac{\partial L}{\partial \Delta u} = p_1 \Delta x + w'(\underline{u})p - p_1(w(\Delta u + \underline{u}) - w(\underline{u})) \\ - p(w'(\Delta u + \underline{u})(1-p) + w'(\underline{u})p) = 0 \tag{A7}$$

$$\mu \geq 0, \mu(U - U_0) = 0. \tag{A8}$$

From (A6), we have $\mu > 0$, since $w'' > 0$ by Lemma 1 and $\Delta u > 0$ by (9). Then, $U = U_0$ using (A8). Moreover, collecting terms in (A7), we get (12), since $p_1 = \frac{1}{\psi_{11}}$ from (9).

Sufficient conditions. Since the constraint is linear in U, necessary conditions are sufficient if V in (4) is concave in $(\Delta u, U)$. We need to verify:

$$\frac{\partial^2 V}{\partial U^2} = -pw''(\Delta u + \underline{u}) - (1-p)w''(\underline{u}) < 0 \tag{A9}$$

$$\frac{\partial^2 V}{\partial \Delta u^2} = p_{11}\big(\Delta x - \big(w(\Delta u + \underline{u}) - w(\underline{u})\big)\big) \\ - p_1\big(w'(\Delta u + \underline{u})(1-p) + w'(\underline{u})p\big) - p_1(1-2p)\big(w'(\Delta u + \underline{u}) - w'(\underline{u})\big) \\ - p(1-p)\big(w''(\Delta u + \underline{u})(1-p) + w''(\underline{u})p\big) < 0 \tag{A10}$$

$$\frac{\partial^2 V}{\partial \Delta u^2}\frac{\partial^2 V}{\partial U^2} - \left(\frac{\partial^2 V}{\partial \Delta u \partial U}\right)^2 \geq 0, \tag{A11}$$

with $\frac{\partial^2 V}{\partial \Delta u \partial U} = -p_1(w'(\Delta u + \underline{u}) - w'(\underline{u})) - p(1-p)(w''(\Delta u + \underline{u}) - w''(\underline{u}))$.

Using Lemma 1, (A9) is indeed negative. By contrast, the sign of (A10) is not warranted. Recall that, $w' > 0, w'' > 0$ from Lemma 1, and $p_1 = \frac{1}{\psi_{11}} > 0$ from Assumption 1, so the second and the fourth terms are negative. However, the first and the third are undetermined. Finally, it is difficult to compute the sign of (A11).

Thus, unfortunately, we cannot be sure that necessary conditions ensure a global maximum.

However, notice that $p_{11} = -\frac{\psi_{111}}{\psi_{11}^3} \leq 0$ from Assumptions 1 and 2, and $\Delta x - \big(w(\Delta u + \underline{u}) - w(\underline{u})\big) > 0$ from (12). Therefore, the first term in (A10) is negative at the solution (12). Moreover, the third term in (A10) is non positive if $p \leq \frac{1}{2}$. It is positive otherwise. Therefore, (A10) is verified, in particular, as long as p is not too much higher than $\frac{1}{2}$, when $1 - 2p < 0$. Therefore, a local maximum is likely to emerge. However, the difficulty of ensuring the sign of (A11) remains.

Appendix A.3. Proofs of Propositions 1 and 5

We omit arguments for the sake of clarity. First, using (19) and (21), we get:

$$U = U_0 - \int_{\underline{\theta}}^{\theta} \psi_2 d\tau.$$

Second, consider the optimal control problem. Let $\Delta u' = y \geq 0$ where y is a control. We associate the adjoint variable η (resp. μ) with U (resp. Δu). The Hamiltonian is:

$$H = \big\{\underline{x} + p\Delta x - w(U + \psi - p\Delta u) - p\big(w(U + \psi + (1-p)\Delta u) - w(U + \psi - p\Delta u)\big)\big\}f \\ - \eta\psi_2 + \mu y.$$

Necessary and transversality conditions. Using (14) and (17), the maximum principle yields[13]:

[13] See Seierstadt and Sydsaeter (1987) [14].

- as necessary conditions:

$$\frac{\partial H}{\partial y} = \mu \leq 0; y\frac{\partial H}{\partial y} = y\mu = 0 \tag{A12}$$

$$\eta' = -\frac{\partial H}{\partial U} = \{w'(\underline{u}) + p(w'(\Delta u + \underline{u}) - w'(\underline{u}))\}f > 0 \tag{A13}$$

$$\mu' = -\frac{\partial H}{\partial \Delta u} = -\{p_1\Delta x + w'(\underline{u})p - p_1(w(\Delta u + \underline{u}) - w(\underline{u})) - p(w'(\Delta u + \underline{u})(1-p) + w'(\underline{u})p)\}f + \eta\psi_{12}p_1, \tag{A14}$$

- as transversality conditions:

$$\eta(\underline{\theta}) \text{ no condition}, \eta(\overline{\theta}) = 0; \tag{A15}$$

$$\mu(\underline{\theta}) = \mu(\overline{\theta}) = 0. \tag{A16}$$

From (A13) and (A15), we have:

$$\eta = -\int_{\theta}^{\overline{\theta}} \{w'(\underline{u}) + p(w'(\Delta u + \underline{u}) - w'(\underline{u}))\}f(\tau)d\tau \leq 0. \tag{A17}$$

Proof of Proposition 1. A fully-separating contract, i.e., $\Delta u' > 0$, implies $y > 0$. Using (A12), we get $\mu = 0$ on Θ. Therefore, μ' is equal to zero on Θ. Using (A14) and (A17), Δu is equal to Δu^* since $p_1 = \frac{1}{\psi_{11}}$ from (14). Using (A12), if it is increasing, i.e., $y^* > 0$, this is the solution.

Proof of Proposition 5. If $y^* < 0$, y must be set equal to zero, and Δu is constant. Consider this occurs on a single interior interval $[\theta_0, \theta_1]$. By continuity and (A16), $\mu(\theta_0)$ and $\mu(\theta_1)$ are both equal to 0. Thus, the constant solution, denoted Δu^k is obtained by integration of (A14) between θ_0 and θ_1, knowing that $\Delta u^*(\theta_0) = \Delta u^*(\theta_1) = \Delta u^k$.

Sufficient conditions. Sufficient conditions require H to be concave in $(\Delta u, U)$. First, it is straightforward to check that $\frac{\partial^2 H}{\partial U^2} = \left(\frac{\partial^2 V}{\partial U^2}\right) f$ and $\frac{\partial^2 H}{\partial \Delta u \partial U} = \left(\frac{\partial^2 V}{\partial \Delta u \partial U}\right) f$.

Second, let us compute $\frac{\partial^2 H}{\partial \Delta u^2}$. Since $p_1 = \frac{1}{\psi_{11}}$ and $p_2 = -\frac{\psi_{12}}{\psi_{11}}$ from (14), we get using (A14):

$$\frac{\partial H}{\partial \Delta u} = (p_1(\Delta x - (w(\Delta u + \underline{u}) - w(\underline{u}))) - p(1-p)(w'(\Delta u + \underline{u}) - w'(\underline{u})))f + \eta p_2. \tag{A18}$$

Therefore, we have:

$$\frac{\partial^2 H}{\partial \Delta u^2} = \left(\frac{\partial^2 V}{\partial \Delta u^2}\right) f + \eta p_{12}.$$

By analogy with the proof of Lemma 4, we cannot ensure the concavity of H. However, it is important to notice that since $f > 0$, V concave implies that H is concave for sure if $\eta p_{12} < 0$. Since $\eta \leq 0$, a sufficient condition is $p_{12} > 0$ (see the discussion below Proposition 4). Therefore, if a global maximum with moral hazard is ensured, a global maximum with the mixed model is.

Appendix A.4. Proof of Proposition 4

Using the concavity of H in Δu, differentiating (A18) and using (19) and (14), we get (29).

References

1. Laffont, J.-J.; Martimort, D. *The Theory of Incentives: The Principal-Agent Model*; Princeton University Press: Princeton, NJ, USA; Oxford, UK, 2002.
2. Bolton, P.; Dewatripont, M. *Contract Theory*; MIT Press: Cambridge, MA, USA, 2005.
3. Kimball, M. Precautionary saving in the small and in the large. *Econometrica* **1990**, *58*, 53–73.

4. Menezes, C.; Geiss, C.; Tressler, J. Increasing downside risk. *Am. Econ. Rev.* **1980**, *70*, 921–932.
5. Eeckhoudt, L.; Gollier, C.; Schneider T. Risk-aversion, prudence and temperance: A unified approach. *Econ. Lett.* **1995**, *48*, 331–336.
6. Crainich, D.; Eeckhoudt, L. La notion économique de prudence, Origine et développements récents. *Rev. Econ.* **2005**, *56*, 1021–1032.
7. Faynzilberg, P.S.; Kumar, P. On the generalized principal-agent problem: Decomposition and existence results. *Rev. Econ. Des.* **2000**, *5*, 23–58.
8. Ollier, S. On the generalized principal agent problem: A comment. *Rev. Econ. Des.* **2007**, *11*, 1–11.
9. Ollier, S.; Thomas, L. Ex post participation constraint in a principal-agent model with adverse selection and moral hazard. *J. Econ. Theory* **2013**, *148*, 2383–2403.
10. Escobar, J.F.; Pulgar, C. Motivating with simple contracts. *Int. J. Ind. Organ.* **2017**, *54*, 192-214.
11. Jullien, B.; Salanié, B.; Salanié, F. Screening risk-averse agents under moral-hazard: single-crossing and the CARA case. *Econ. Theory* **2007**, *30*, 151–169.
12. Laffont, J.-J.; Tirole, J. Using cost observation to regulate firms. *J. Political Econ.* **1986**, *94*, 614–641.
13. Myerson, R. Optimal coordination mechanisms in generalized principal-agent problems. *J. Math. Econ.* **1982**, *10*, 67–81.
14. Seierstad, A.; Sydsæter, K. *Optimal Control Theory with Economic Applications*; North-Holland: Amsterdam, The Netherlands, 1987.