

# Risk-sharing or risk-taking? An incentive theory of counterparty risk, clearing and margins\*

Bruno Biais<sup>†</sup>      Florian Heider<sup>‡</sup>      Marie Hoerova<sup>§</sup>

March 18, 2014

## Abstract

Derivatives trading, motivated by risk-sharing, can breed risk-taking. Bad news about the hedged risk increases the expected liability of the protection seller, undermining her risk-prevention incentives. This creates endogenous counterparty risk and contagion from news about the hedged risk to the balance sheet of the protection seller. Margin calls after bad news can improve protection sellers' incentives and enhance the ability to share risk. Central clearing can provide insurance against counterparty risk but must be designed to preserve risk-prevention incentives.

JEL CLASSIFICATION: G21, G22, D82

KEYWORDS: Hedging; Insurance; Derivatives; Moral hazard; Risk management; Counterparty risk; Contagion; Central clearing; Margin requirements

March 18, 2014

---

\*We would like to thank the Editor (Cam Harvey), an Associate Editor and two anonymous referees, our discussants Ulf Axelson, Jonathan Berk, Sugato Bhattacharyya, Bruce Carlin, Simon Gervais, Artashes Karapetyan, Lauri Vilmi as well as numerous seminar and conference participants for their comments and suggestions. A previous version of the paper was circulated under the title “Risk-sharing or risk-taking? Counterparty risk, incentives and margins”. The views expressed do not necessarily reflect those of the European Central Bank or the Eurosystem. Biais gratefully acknowledges the support of the European Research Council.

<sup>†</sup>Toulouse School of Economics (CNRS-CRM, Fédération des Banques Françaises Chair on the Investment Banking and Financial Markets Value Chain at IDEI), email: bruno.biais@univ-tlse1.fr

<sup>‡</sup>European Central Bank, Financial Research Division, email: florian.heider@ecb.int

<sup>§</sup>European Central Bank, Financial Research Division, email: marie.hoerova@ecb.int

# 1 Introduction

Many financial institutions trade derivatives in order to hedge risk. Such insurance, however, can be effective only if their counterparties do not default. Optimal clearing mechanisms could reduce that risk. Margin calls can increase the amount of resources available to make payments. Centralized clearing enables investors to mutualize counterparty risk. Indeed, regulators and policy makers have mandated centralized clearing of risky financial instruments in the US (with the Dodd–Frank Wall Street Reform Act) as well as in Europe (with the European Market Infrastructure Regulation.) There is considerable debate, however, about the optimal design of market infrastructures (see, e.g., Bernanke (2011) and Roe (2013)).

This paper studies how clearing should be designed to enable optimal risk–sharing via derivatives trading. Our approach is based on the observation that derivatives trading and clearing do not simply reallocate existing risk, they also affect the incentives of market participants to control risk or engage in risk–taking. Clearing mechanisms should therefore be designed subject to incentive compatibility constraints. Otherwise, derivatives trading could lead protection sellers to engage in excessive risk–taking, and centralized clearing could increase (instead of reducing) aggregate risk. Thus, we take an optimal contracting approach to the design of derivative contracts and clearing mechanisms.

Our model features a population of protection buyers, a population of protection sellers, and a central counterparty (hereafter CCP). The protection buyers’ assets (e.g., corporate or real-estate loans held by commercial banks) are exposed to risk. Their portfolios may be diversified across industries and geographic locations, but are still exposed to aggregate risk, e.g., from nationwide business cycles or bubbles. Due to leverage or regulatory constraints, such as risk–weighted capital requirements, protection buyers would benefit from hedging that risk. To do so, they turn to protection sellers, such as investment banks or insurance companies selling Credit Default Swaps (CDS).<sup>1</sup> Protection sellers, however, can make insurance payments only if their assets are sufficiently valuable. Hedging, therefore, can be effective only if protection sellers are incentivized to maintain the value of their assets.

---

<sup>1</sup>A prominent example was AIG. 72% of the CDS it had sold by December 2007 were used by banks for capital relief (European Central Bank, 2009).

Otherwise, protection buyers are exposed to default risk.

To reduce downside risk on their assets, protection sellers must exert effort, e.g., to screen out bad loans or monitor borrowers. Given the complex and opaque nature of the protection sellers' balance-sheets and the risks embedded therein, their risk-prevention effort is difficult to observe and monitor for outside parties. At the same time, sound risk-management is costly because it requires time and resources. One way to shirk on risk-management effort (and reduce its cost) is to rely on easily available but superficial information such as ready-made ratings instead of performing an own investigation of risk. Another way is to finance long-term assets with short-term debt without taking into account the risk involved when rolling over the debt. Thus, we consider a setting with unobservable but costly effort, i.e., moral hazard.<sup>2</sup>

The first contribution of this paper is to show that derivatives trading can undermine the incentives of protection sellers and therefore create counterparty risk and contagion. Ex-ante, when the protection seller enters the position, the derivative contract is neither an asset nor a liability. For example, the seller of a CDS pays the buyer in case of credit events (default, restructuring) but collects an insurance premium otherwise, and on average these costs and benefits offset each other. But, upon observing a strong drop in real estate prices, sellers of subprime-mortgage CDS anticipate to be liable for insurance payments. This liability undermines their incentives to exert risk-prevention effort. Similar to the debt-overhang effect analyzed by Myers (1977), the protection seller bears the full cost of such effort while part of its benefits accrue to the protection buyer.<sup>3</sup> It can then happen, in equilibrium, that the protection seller stops exerting risk-prevention effort on his assets, which raises counterparty default risk.<sup>4</sup> Our analysis thus identifies a channel through which derivatives trading can propagate risk. For simplicity, in our model, without moral hazard, the riskiness of the assets of the protection sellers is independent from that of the protection buyers assets.

---

<sup>2</sup>In the present analysis, the unobservable action of the agent affects the cash-flows in the sense of first-order stochastic dominance, as in Holmström and Tirole (1998). In a previous version of the paper, we showed that qualitatively identical results hold when the unobservable action leads to an increase in risk in the sense of second-order stochastic dominance (risk-shifting), in the spirit of Jensen and Meckling (1976).

<sup>3</sup>Note however that instead of exogenous debt as in Myers (1977) our model involves endogenous liabilities pinned down in an optimal hedging contract.

<sup>4</sup>For example, Lehman Brothers and Bear Stearns defaulted on their CDS derivative obligations because of losses incurred on their other investments, in particular sub-prime mortgages.

With moral hazard, in contrast, bad news about the risk of the protection buyers reduce protection sellers' risk-prevention effort, and increase downside risk for their assets. This generates contagion (endogenous correlation) between the two asset classes.<sup>5</sup>

The second contribution of this paper is to analyze how margins and clearing affect the effectiveness of hedging and the possibility of contagion. The optimal CCP stipulates not only the transfers to (or from) protection buyers and sellers, but also the circumstances under which protection sellers must liquidate a fraction of their risky assets and deposit the resulting cash on a margin account. The cost of such liquidation is the wedge between what the assets could have earned and the lower return on cash in a margin deposit. On the other hand, the cash in the margin account is no longer under the control of the protection seller, and therefore is ring-fenced from moral hazard. We show that it is optimal to call margins (only) after bad news about the asset underlying the derivatives trade. After such bad news, the derivative position of the protection sellers is a liability for them, reducing their incentives to exert risk-prevention effort. It is under those circumstances that the incentives of the protection sellers need to be maintained. This is achieved by liquidating some of the protection sellers' assets and depositing them in the margin account, which reduces their temptation to shirk. Thus, variation margins relax incentive constraints and therefore increase the ability to offer insurance without creating counterparty risk.

While this benefit of margins could be reaped with bilateral clearing, risk-sharing can be further improved with centralized clearing. In particular, CCPs enable market participants to mutualize counterparty default risk. Note however that insurance against counterparty risk can generate additional moral hazard issues. Since margin deposits are costly, market participants are reluctant to make them, especially when they are insured against counterparty risk by the CCP. Thus our analysis implies that margin calls in CCPs should be mandatory, rather than determined bilaterally. Our theory also implies that financial institutions with lower pledgeable income should make larger margin deposits. Lower pledgeability can arise due to insufficient equity capital, weak risk-management (see Ellul and Yerramili, 2010), or complex and opaque activities. Thus there is substitutability between i) margins and ii) equity capital, effective risk-management and transparency.

---

<sup>5</sup>This incentive-based theory of contagion differs from the analyses of systemic risk offered by Freixas, Parigi and Rochet (2000), Cifuentes, Shin and Ferrucci (2005), and Allen and Carletti (2006).

## MAYBE NEED TO SAY A BIT MORE ON OUR RESULTS AND IMPLICATIONS

WE NEED TO SAY MORE ABOUT THE DIFFERENCE OF OUR MODEL TO HOLMSTROM AND TIROLE. Our model delivers full risk-sharing conditional on the signal, i.e., our contract is ex-post Pareto efficient (both after the signal and after the effort decision). Therefore, there is no scope for re-negotiation ex post. This is different from the standard HT model and different from the debt overhang re-negotiation. Also, in our model, the risk is borne by the risk-neutral party (in standard models, it is the risk-averse agent that does the unobservable effort and one has to expose him to risks to incentivize him and this is inefficient).

Our paper is related to the literature on financial risk insurance, on margins and clearing, and on liquidation and collateral.

Thompson (2010) assumes moral hazard on the part of the protection seller. In his model, however, i) the protection buyer is privately informed about his own risk and ii) his hidden action affects the liquidity of the assets she invests in. In this context, moral hazard alleviates adverse selection and therefore *enhances* the provision of insurances. This is very different from our analysis, where there is no adverse selection and moral hazard *impedes* the provision of insurance. Allen and Carletti (2006) and Parlour and Plantin (2008) analyze credit risk transfer in banking. Again, their analyses are very different from ours, since the friction in Allen and Carletti (2006) is cash-in-the-market pricing for long-term assets, while in Parlour and Plantin (2008) it is moral hazard problem on the side of the insured. Bolton and Oehmke (2013) borrow from our framework the mechanism by which posting collateral or margins deters risk-taking. They use it to address another issue than the issue on which we focus. They show that effective seniority for derivatives transfers to the firm's debtholders credit risk that could be borne more efficiently by the derivative market. Both Stephens and Thompson (2011) and our paper analyze how market imperfection raise counterparty risk. In their heterogeneous types model, increased competition leads to lower insurance premia and riskier protection sellers' types. In contrast, in our homogeneous types model, for any given level of competition, bad news can undermine protection sellers' risk prevention incentives.

Acharya and Bisin (2011) study the inefficiency which can arise between one protection

seller and several protection buyers. In an over-the-counter market no buyer can control the trades of the seller with other buyers. Yet, when the protection seller contracts with an additional protection buyer, this exerts a negative externality on the other protection buyers since it increases their counterparty risk (see also Parlour and Rajan, 2001). Acharya and Bisin (2011) show how, with centralized clearing and trading, such externality can be avoided, by implementing price schedules penalizing the creation of counterparty risk. In contrast with the excessive positions analyzed by Acharya and Bisin (2011), the hidden action generating a moral hazard problem in our model cannot be observed, and therefore cannot be penalized with centralized clearing. Thus, while in Acharya and Bisin (2011) optimality entails conditioning *prices* on all trades, in the present paper it entails constraints on the *quantities* of insurance and assets under management.

Margins can be understood as collateral deposited by the agent to reduce the risk of the principal. However, our focus on hedging and derivatives differs from that of papers studying borrowing and lending (see, e.g, Bolton and Sharfstein (1990), Holmstrom and Tirole (1998), Acharya and Viswanathan (2011)). In our analysis, as in derivative markets, the margin is called *before* effort decisions are taken and output is realized. In contrast, collateral in a financing context is liquidated *after* effort has been exerted and output realized (as, e.g., in Holmstrom and Tirole (1998), Tirole (2006) and Acharya and Viswanathan (2011)). Margins in derivative contracts differ from collateral in loan contracts precisely because they are called before maturity or default. Our paper offers the first analysis of this essential feature of margins and its incentive properties.

The model is presented in the Section 3, which also analyzes the benchmark case in which effort is observable. Section 4 analyzes optimal contracting under moral hazard. Section 5 concludes. Proofs are in the Appendix.

## 2 Model and First–Best Benchmark

### 2.1 The model

There are three dates,  $t = 0, 1, 2$ , a mass–one continuum of protection buyers, a mass–one continuum of protection sellers and a Central Clearing Platform, hereafter referred to as the

CCP. At  $t = 0$ , the parties design and enter the contract. At  $t = 1$ , investment decisions are made. At  $t = 2$ , payoffs are received.

**Players and assets.** Protection buyers are identical, with twice differentiable concave utility function  $u$ , and are endowed with one unit of an asset with random return  $\tilde{\theta}$  at  $t = 2$ . For simplicity, we assume  $\tilde{\theta}$  can only take on two values:  $\bar{\theta}$  with probability  $\pi$  and  $\underline{\theta}$  with probability  $1 - \pi$ , and we denote  $\Delta\theta = \bar{\theta} - \underline{\theta}$ . The risk  $\tilde{\theta}$  is the same for all protection buyers.<sup>6</sup>

Protection buyers seek insurance against the risk  $\tilde{\theta}$  from protection sellers who are risk-neutral and have limited liability. Each protection seller  $j$  has an initial amount of cash  $A$ . At time  $t = 1$ , this initial balance sheet can be split between two types of assets: i) low risk, low return assets such as Treasuries (with return normalized to 1), and ii) risky assets returning  $\tilde{R}_j$  per unit at  $t = 2$ . The protection seller has unique skills (unavailable to the protection buyer or the CCP) to manage the risky assets and earn excess return. After this initial investment allocation decision, the protection seller makes a risk-management decision at  $t = 1$ . To model risk-management in the simplest possible way, we assume that each seller  $j$  can undertake a costly effort to make her assets safer. If she undertakes the effort, the per unit return  $\tilde{R}_j$  is  $R$  with probability one. If she does not exert the effort, then the return is  $R$  with probability  $p < 1$  and zero with probability  $1 - p$ . The risk-management process reflects the unique skills of the protection seller and is therefore difficult to observe and monitor by outside parties. This opacity gives rise to moral hazard which we model by assuming that the risk management effort decision is unobservable.

Exerting the effort costs  $C$  per unit of assets under management at  $t = 1$  (we explain below why there could be less assets under management at  $t = 1$  than initial assets in place at  $t = 0$ ).<sup>7</sup> Because protection seller assets are riskier without costly effort, we also call the

---

<sup>6</sup>At the cost of unnecessarily complicating the analysis, we could also assume that the risk has an idiosyncratic component. This component would not be important as protection buyers could hedge this risk among themselves, without seeking insurance from protection sellers.

<sup>7</sup>We show later that our results are unchanged when we allow the unit cost  $C$  to increase (linearly) with assets under management, which makes the overall cost of risk-management effort convex.

decision not to exert effort “risk-taking”.<sup>8</sup> Undertaking risk-management effort is efficient,

$$R - C > pR, \quad (1)$$

i.e., the expected net return is larger with risk-management effort than without it. We also assume that when protection seller exerts risk management effort, return on his assets is higher than the return on the safe asset,

$$R - C > 1, \quad (2)$$

For simplicity, once protection sellers have decided whether or not to undertake risk-management effort, the risk of their assets  $\tilde{R}_j$  is independent across sellers and, moreover, it is independent of protection buyers’ risk  $\tilde{\theta}$ . To allow protection sellers that exert effort to fully insure buyers, we assume  $AR \geq \pi\Delta\theta$ .

**Advance information.** At the beginning of  $t = 1$ , before investment and risk management decisions are made, a public signal  $\tilde{s}$  about protection buyers’ risk  $\tilde{\theta}$  is observed. For example, when  $\tilde{\theta}$  is the credit risk of real estate portfolios,  $\tilde{s}$  should be seen as real estate price index. Denote the conditional probability of a correct signal with

$$\lambda = \text{prob}[\bar{s}|\bar{\theta}] = \text{prob}[\underline{s}|\underline{\theta}].$$

The probability  $\pi$  of a good outcome  $\bar{\theta}$  for the protection buyer’s risk is updated to  $\bar{\pi}$  upon observing a good signal  $\bar{s}$  and to  $\underline{\pi}$  upon observing a bad signal  $\underline{s}$ , where, by Bayes’ law,

$$\bar{\pi} = \text{prob}[\bar{\theta}|\bar{s}] = \frac{\lambda\pi}{\lambda\pi + (1-\lambda)(1-\pi)} \quad \text{and} \quad \underline{\pi} = \text{prob}[\underline{\theta}|\underline{s}] = \frac{(1-\lambda)\pi}{(1-\lambda)\pi + \lambda(1-\pi)}.$$

We assume that  $\lambda \geq \frac{1}{2}$ . If  $\lambda = \frac{1}{2}$ , then  $\bar{\pi} = \pi = \underline{\pi}$  and the signal is completely uninformative. If  $\lambda > \frac{1}{2}$ , then  $\bar{\pi} > \pi > \underline{\pi}$ , i.e., observing a good signal  $\bar{s}$  increases the probability of a good outcome  $\bar{\theta}$  whereas observing a bad signal  $\underline{s}$  decreases the probability of a good outcome  $\bar{\theta}$ . If  $\lambda = 1$ , then the signal is perfectly informative.

### Central counterparty, contracts and margins.

---

<sup>8</sup>Here risk-management effort makes protection sellers’ assets safer in the sense of first-order stochastic dominance. In an earlier version of the paper we show that our results are robust when risk-management effort improves risk in the sense of mean-preserving spreads.



In practice, protection buyers and protection sellers contract bilaterally, and the CCP then interposes between contracting parties. Thus, the contract between the protection buyer and protection seller is transformed into two contracts, one between the seller and the CCP and another one between the buyer and the CCP (a process called novation). In our model, for simplicity, we by-pass the first step (bilateral contracting), and analyze directly the contracts between the CCP and protection buyers and sellers. This enables us to approach the problem from a mechanism design viewpoint in which the CCP designs an optimal mechanism for buyers and sellers.

Correspondingly, the CCP is modeled as a public utility designed to maximize the welfare of its members (i.e., it acts as the social planner). For simplicity, we assume the CCP maximizes expected utility of protection buyers subject to the participation constraint of the protection sellers.<sup>9</sup>

At  $t = 0$ , the CCP specifies transfers  $\tau^S$  between protection sellers and the CCP at  $t = 2$ , and transfers  $\tau^B$  between protection buyers and the CCP at  $t = 2$ . Positive transfers  $\tau^S, \tau^B > 0$  represent payments from the CCP to sellers and buyers, while negative transfers represent payments from sellers and buyers to the CCP. The transfers  $\tau^S$  and  $\tau^B$  at  $t = 2$  are contingent on all available information at that time. This information consists of the buyers' risk  $\tilde{\theta}$ , the signal  $\tilde{s}$  and the set of all the protection sellers' asset returns  $\tilde{R}$ . Hence, we write  $\tau^S(\tilde{\theta}, \tilde{s}, \tilde{R})$  and  $\tau^B(\tilde{\theta}, \tilde{s}, \tilde{R})$ . Since the transfers are contingent on final asset values as well as advance public information about those values (that could be conveyed, e.g., by asset prices), we can think of them as transfers specified by derivative contracts.

The transfers between the CCP and its members reflect the initial underlying bilateral contract, which is novated, and mutualization across all bilateral contracts. Hence, the transfers depend not only on a protection seller individual asset return  $\tilde{R}_j$ , as would be the case in a bilateral contract without the CCP, but depend on all sellers' asset returns  $\tilde{R}$ . This is because the latter affect the amount of resources available to the CCP to insure its members against counterparty risk.

Figure 1 illustrates how the CCP sits in between protection buyers and sellers.

---

<sup>9</sup>While this is only one point on the Pareto frontier, in the first-best all other Pareto optima would entail the same real decisions, i.e., the same risk-sharing and productive efficiency. In the second-best, changing the bargaining would change the structure of the risk sharing, without altering our qualitative results.

### Insert Figure 1 here

The contract between the CCP and its members not only specifies transfers, it can also request margin deposits. Because the CCP has no ability to manage risky, opaque assets, it only accepts as margin deposits safe, transparent ones, such as cash or Treasuries that are not subject to information asymmetry problems.<sup>10</sup> One can therefore interpret margins as an institutional arrangement that affects the split of the seller's balance sheet between transparent assets and opaque investments. Margins “ring-fence” a fraction of the protection sellers' assets from moral-hazard. However, margins incur the opportunity cost of foregoing the excess return of the risky asset,  $R - C - 1$ . The margin can be contingent on all information available at time 1, i.e., the signal  $\tilde{s}$ . We denote the fraction of the protection seller's balance sheet deposited on the margin account by  $\alpha(\tilde{s})$ .

The CCP is subject to budget-balance, or feasibility, constraints at  $t = 2$ . For each joint realization of buyers' risk  $\tilde{\theta}$ , the signal  $\tilde{s}$  and sellers' asset returns  $\tilde{R}$ , aggregate transfers to protection buyers cannot exceed aggregate transfers from protection sellers (the CCP has no resources of its own):

$$\tau^B(\theta, s, R) \leq -\tau^S(\theta, s, R), \quad \forall(\theta, s, R). \quad (3)$$

Transfers from protection sellers are constrained by limited liability,

$$-\tau^S(\theta, s, R) \leq \alpha(s)A + (1 - \alpha(s))AR, \quad \forall(\theta, s, R). \quad (4)$$

A protection seller cannot make transfers larger than what is returned by the fraction  $(1 - \alpha(s))$  of assets under her management and by the fraction  $\alpha(s)$  of assets she deposited on the margin account. Finally, the fraction of assets deposited must be between zero and one,

$$\alpha(s) \in [0, 1] \quad \forall s. \quad (5)$$

The sequence of events is summarized in Figure 2.

### Insert Figure 2 here

---

<sup>10</sup>That assets with low information sensitivity are used as collateral is in line with Gorton and Pennacchi (1990).

## 2.2 First-best: observable effort

In this subsection we consider the case in which protection sellers' risk-management effort is observable, so that there is no moral hazard and the first-best is achieved. While implausible, this case offers a benchmark against which we will identify the inefficiencies arising when protection seller's risk-management effort is not observable.

Protection sellers are requested to exert risk-management effort when offering protection since doing so increases the resources available for risk-sharing (see (1)). Margins are not used since they are costly (see (2)) and offer no benefit when risk-management effort is observable. The CCP chooses transfers to buyers and sellers,  $\tau^B(\tilde{\theta}, \tilde{s}, \tilde{R})$  and  $\tau^S(\tilde{\theta}, \tilde{s}, \tilde{R})$ , to maximize buyers' utility

$$E[u(\tilde{\theta} + \tau^B(\tilde{\theta}, \tilde{s}, \tilde{R}))] \quad (6)$$

subject to the feasibility (3) and limited liability (4) constraints, as well as the constraint that protection sellers participate and join the CCP. By joining (and exerting effort), sellers obtain  $E[\tau^S(\tilde{\theta}, \tilde{s}, \tilde{R})] + A(R - C)$ . If they do not join and thus do not sell protection, they obtain  $A(R - C)$ .<sup>11</sup> The protection sellers' participation constraint in the first-best therefore is

$$E[\tau^S(\tilde{\theta}, \tilde{s}, \tilde{R})] \geq 0. \quad (7)$$

Proposition 1 states the first-best outcome. Since protection sellers exert risk-management effort, the return  $\tilde{R}$  is always equal to  $R$  and we drop the reference to the return in the transfers  $\tau^B$  and  $\tau^S$  for ease of notation.

**Proposition 1** *When effort is observable, the optimal contract entails effort, provides full insurance, is actuarially fair and does not react to the signal. Margins are not used. The transfers are given by*

$$\begin{aligned} \tau^B(\bar{\theta}, \bar{s}) &= \tau^B(\bar{\theta}, \underline{s}) = E[\tilde{\theta}] - \bar{\theta} < 0 \\ \tau^B(\underline{\theta}, \bar{s}) &= \tau^B(\underline{\theta}, \underline{s}) = E[\tilde{\theta}] - \underline{\theta} > 0 \\ \tau^B(\theta, s) &= -\tau^S(\theta, s), \forall(\theta, s) \end{aligned}$$

---

<sup>11</sup>Without derivative trading, protection sellers always exert effort since it is efficient to do so (see condition (2)).

The first-best contract fully insures protection buyers. Their marginal utility, and hence their consumption, is the same across all realizations of their risky asset  $\theta$  and the signal  $s$ . The transfers are independent of the signal and ensure a consumption level equal to the expected value of the risky asset,  $E[\tilde{\theta}]$ . The first-best insurance contract is actuarially fair since the expected transfer from protection sellers to protection buyers is zero,  $E[\tau^B(\tilde{\theta}, \tilde{s})] = -E[\tau^S(\tilde{\theta}, \tilde{s})] = 0$ . We assume

$$AR > \pi\Delta\theta, \tag{8}$$

so that, in the first-best, the aggregate resources of the protection sellers are large enough to fully insure the protection buyers.

In our simple model, when effort is observable, each transfer to a protection buyer  $\tau^B$  is matched by an opposite transfer from a protection seller and margins are not needed. Thus the contract can be implemented bilaterally and the CCP is not needed. Of course, this reflects our simplifying assumption that, under effort,  $R$  is obtained for sure. If protection sellers could default, even with high effort, the CCP would be useful, in the first best, to mutualize default risk. As shown in the next sections, even in the simple case where effort precludes default, with moral hazard, the CCP plays a useful role.

The first best transfers,  $\tau^B(\theta, s)$  and  $\tau^S(\theta, s)$ , can be implemented with forward contracts. The protection buyer sells the underlying asset forward, at price  $F = E[\tilde{\theta}]$ . When the final value of the asset is worth  $\bar{\theta}$ , the protection buyer must deliver at the relatively low forward price  $F$ . But, when the final value of the asset is low  $\underline{\theta}$ , the forward price is relatively high. This provides insurance to the protection buyer.

While we only consider transfers at  $t = 2$ , and not explicitly at  $t = 1$ , this is without loss of generality, because any other trading arrangement can be replicated with transfers at  $t = 2$  and margins. Consider for example spot trading in which at  $t = 1$ , before the realization of the signal, the protection seller uses some of his initial assets  $A$  to acquire the protection buyers' asset at price  $S$ . Because there is no discounting, this is equivalent for the protection buyer to a constant transfer  $S$  at time 2. This can be achieved within the mechanism we analyze, by depositing  $S$  on the margin account at  $t = 1$  and letting  $\tau^B(\theta, s) = S$ , irrespective of the realization of  $\theta$  and  $s$ . Proposition 1 shows, however, that this is dominated by forward trading. Forward trading is more efficient, because it makes it

possible to keep the assets under the management of the protection seller until  $t = 2$  and earn a larger return  $(R - C)$  than when investing in the risk free asset.

### 3 Protection-seller moral-hazard

In the previous section, we examined the hypothetical case in which protection sellers' risk-management effort is observable and can therefore be requested by protection buyers. We now move on to the more realistic situation in which risk-management effort is not observable and there is moral-hazard on the side of protection sellers.

If protection buyers want protection sellers to exert risk-management effort, then it must be in sellers' own interest to do so after observing the signal  $s$  about buyers' risk  $\tilde{\theta}$ . The incentive compatibility constraint under which a protection seller exerts effort after observing  $s$  is:

$$\begin{aligned} E[\tau^S(\tilde{\theta}, \tilde{s}, \tilde{R}) + \alpha(\tilde{s})A + (1 - \alpha(\tilde{s}))A(\tilde{R} - C) | e = 1, \tilde{s} = s] \\ \geq E[\tau^S(\tilde{\theta}, \tilde{s}, \tilde{R}) + \alpha(\tilde{s})A + (1 - \alpha(\tilde{s}))A\tilde{R} | e = 0, \tilde{s} = s]. \end{aligned}$$

The left-hand side is a protection seller's expected payoff if she exerts risk-management effort. The effort costs  $C$  per unit of assets she still controls,  $(1 - \alpha(s))A$ . The right-hand side is her (out-of-equilibrium) expected payoff if she does not exert effort and therefore does not incur the cost  $C$ .

Without effort, her assets under management return  $R$  with probability  $p$  and zero with probability  $1 - p$ . In order to relax the incentive constraint, the CCP requests the largest possible transfer from a protection seller when  $\tilde{R} = 0$ :  $-\tau^S(\tilde{\theta}, \tilde{s}, 0) = \alpha(\tilde{s})A$ . This rationalizes the stylized fact that, in case of default of the protection seller, the CCP seizes her deposits and uses them to pay protection buyers.

With effort, the protection seller's assets under management are safe, with  $\tilde{R} = R$ . For brevity, we write  $\tau^S(\tilde{\theta}, \tilde{s}, R)$  as  $\tau^S(\tilde{\theta}, \tilde{s})$ . The incentive constraint after observing  $s$  is then

$$\begin{aligned} E[\tau^S(\tilde{\theta}, \tilde{s}) | \tilde{s} = s] + \alpha(s)A + (1 - \alpha(s))A(R - C) \\ \geq p \left( E[\tau^S(\tilde{\theta}, \tilde{s}) | \tilde{s} = s] + \alpha(s)A + (1 - \alpha(s))AR \right), \end{aligned}$$

or, using the notion of “pledgeable return”  $\mathcal{P}$  (see Tirole, 2006),

$$\mathcal{P} \equiv R - \frac{C}{1-p}, \quad (9)$$

the incentive compatibility constraint rewrites as

$$\alpha(s)A + (1 - \alpha(s))A\mathcal{P} \geq E[-\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = s]. \quad (10)$$

The right-hand side is what protection sellers expect to pay to the CCP after seeing the signal about buyers’ risk. The left-hand side is the amount that protection sellers’ can pay (or pledge) to the CCP without undermining their incentive to exert risk-management effort. The left-hand side is positive since the assumption that effort is efficient, condition (1), ensures positive pledgeable return,  $\mathcal{P} > 0$ . The right-hand side is positive when, conditional on the signal, a protection seller expects, on average, to make transfers to the CCP. If after seeing the signal she expects, on average, to receive transfers from the CCP, then the right-hand side is negative and the incentive constraint does not bind. This is an important observation to which we return later.

When the pledgeable return  $\mathcal{P}$  is sufficiently high, then protection sellers’ incentive problem does not matter because the first-best allocation (stated in Proposition 1) satisfies the incentive-compatibility constraint (10) after any signal. The exact condition is given in the following lemma.

**Lemma 1** *When risk-management effort is not observable, the first-best can be achieved if and only if the pledgeable return on assets is high enough:*

$$A\mathcal{P} \geq (\pi - \underline{\pi})\Delta\theta = E[\tilde{\theta}] - E[\tilde{\theta}|\tilde{s} = \underline{s}]. \quad (11)$$

The threshold for the pledgeable return on assets, beyond which full risk-sharing is possible despite protection seller moral-hazard, is given by the difference between the unconditional expectation of buyers’ risk  $\tilde{\theta}$  and the conditional expectation of this risk after a low signal (indicating a bad outcome is more likely). The threshold increases, making it more difficult to attain the first-best, when buyers’ assets are riskier (larger  $\Delta\theta$ ) and, interestingly, when there is *better* information about this risk (larger  $\lambda$  leading to a lower  $\underline{\pi}$ ). Thus, Lemma 1 has the following corollary.

**Corollary 1** *When the signal is uninformative,  $\lambda = \frac{1}{2}$ , the first-best is always reached since  $(\pi - \underline{\pi})\Delta\theta = 0$ .*

In what follows, we focus on the case in which protection seller moral-hazard matters and full insurance is not feasible, as (11) does not hold.

### 3.1 Effort after both signals

In this section, we study the contract providing the protection seller the incentives to exert risk-management effort both after positive and after negative signals. While margins were not useful without moral-hazard (as discussed in Subsection 2.2), they may be useful now. When a protection seller exerts risk-management effort after both signals, her participation constraint is

$$E[\alpha(\tilde{s})A + (1 - \alpha(\tilde{s}))A(\tilde{R} - C) + \tau^S(\tilde{\theta}, \tilde{s}, \tilde{R})|e = 1] \geq A(R - C).$$

Since, on the equilibrium path, the protection sellers exert effort, we have  $\tilde{R} = R$  and again, for brevity, we write the transfer to a protection seller as  $\tau^S(\tilde{\theta}, \tilde{s})$ . Collecting terms, the participation constraint is

$$E[\tau^S(\tilde{\theta}, \tilde{s})] \geq E[\alpha(\tilde{s})]A(R - C - 1), \tag{12}$$

The expected transfers from the CCP to a protection seller (left-hand-side) must be high enough to compensate her for the opportunity cost of the expected use of margins (right-hand-side). Thus, if margins are used, the contract is not actuarially fair.

The CCP chooses transfers to protection buyers  $\tau^B(\tilde{\theta}, \tilde{s})$  and protection sellers,  $\tau^S(\tilde{\theta}, \tilde{s})$ , as well as margins  $\alpha(\tilde{s})$ , to maximize buyers' utility (6) subject to the feasibility constraints (3), the constraint that the fraction  $\alpha$  be in  $[0, 1]$  (5), and the incentive (10), limited liability (4), and participation (12) constraints.

The next proposition collects first results on how resources are optimally transferred between protection sellers and protection buyers.

**Proposition 2** *In the optimal contract with risk-management effort, the feasibility constraints (3) bind for all  $(\theta, s)$ , the limited liability constraints (4) are slack in state  $(\bar{\theta}, s)$  for each  $s$ , and the participation constraint (12) binds.*

Protection sellers earn no rents and all resources available for insurance are passed on to protection buyers. Protection sellers' limited liability is not an issue when the value of the protection buyers' asset is  $\bar{\theta}$ , since in that state risk-sharing implies positive transfers to protection sellers.

Using the binding feasibility constraints, we can rewrite the incentive constraint (10) as

$$\alpha(s)A + (1 - \alpha(s))A\mathcal{P} \geq E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = s] \quad (13)$$

Incentive compatibility implies that the expected transfers to the protection buyer be no larger than the sum of the returns on i) the assets deposited on the margin account and on ii) those left under the protection seller's management. The pledgeable return on assets under management is smaller than the physical net return,  $\mathcal{P} < R - C$ , because there is moral hazard when exerting effort to manage the risk of those assets. The pledgeable return on assets deposited on the margin account is equal to their physical return of one since they are "ring-fenced" from moral-hazard in risk-management. When the moral hazard is severe,  $\mathcal{P} < 1$ , then depositing assets on the margin account relaxes the incentive constraint and thus allows for higher transfers to protection buyers. This is the benefit of margins. But assets deposited on the margin account incur an opportunity cost  $R - C - 1$  to protection sellers. This basic tradeoff leads to the following proposition:

**Proposition 3** *In the optimal contract with risk-management effort, margins are not used after  $s$  if the incentive constraint given  $s$  is slack or if the moral-hazard is not severe, i.e.,  $\mathcal{P} \geq 1$ .*

When the incentive constraint after  $s$  is slack, then depositing assets on the margin account offers no incentive benefit and only incurs the opportunity cost. When the pledgeable return of assets under management (weakly) exceeds the pledgeable return of assets deposited on the margin account, then margins also do not offer any incentive benefit since they actually tighten the incentive constraint.

To keep the next steps of the analysis tractable, we make the following simplifying assumption:

$$AR > \bar{\pi}\Delta\theta - \frac{\text{prob}[s]}{\text{prob}[\bar{s}]}A\mathcal{P}, \quad (14)$$



The assumption guarantees, as we will show, a slack limited liability constraint for transfers from a protection seller to the CCP when there is a good signal,  $\bar{s}$ , but buyers' asset return is low,  $\underline{\theta}$ . We discuss this assumption in more detail once we have solved for the optimal transfers,  $\tau^B(\tilde{\theta}, \tilde{s})$  and  $\tau^S(\tilde{\theta}, \tilde{s})$ . Given (14) and Proposition 2, we only need to consider the limited liability constraint in state  $(\underline{\theta}, \underline{s})$ .

The next proposition states that moral-hazard problem matters only after a bad signal.

**Proposition 4** *In the optimal contract with risk-management effort, the incentive constraint (13) binds after a bad signal, but is slack after a good signal. Hence there is no margin call after a good signal, i.e.,  $\alpha(\bar{s}) = 0$ .*

After observing a bad signal about the underlying risk, a protection seller's position is a liability to her,  $E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] < 0$ . This undermines her incentives to exert risk-management effort. She has to bear the full cost of effort while the benefit of staying solvent accrues in part to protection buyers in the form of the (likely) transfer to the CCP. This is in line with the debt-overhang effect (Myers, 1977).

In contrast, there is no moral-hazard problem for a protection seller after observing a good signal. A good signal indicates that her position is an asset at this point of time,  $E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}] > 0$ . This strengthens her incentives to exert risk-management effort. In a sense, after a good signal, since the protection seller's position has become an asset for her, it increases the income she can pledge. In contrast, the loss she expects after a bad signal reduces her pledgeable income.

We are now ready to characterize the optimal contract between the CCP, protection buyers and protections seller that exert risk management effort. It is convenient to first characterize optimal transfers as a function of the margin after a bad signal,  $\alpha(\underline{s})$ , and later examine the optimal margin call after a bad signal. Expected transfers conditional on the signal (as a function of  $\alpha(\underline{s})$ ) are given by the binding participation constraint (Proposition 2) and the incentive constraint after a bad signal (Proposition 4),

$$E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] = A[\alpha(\underline{s}) + (1 - \alpha(\underline{s}))\mathcal{P}] \quad (15)$$

$$E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}] = -\frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A[\alpha(\underline{s})(R - C) + (1 - \alpha(\underline{s}))\mathcal{P}] \quad (16)$$

The next proposition characterizes the transfers in each possible state:

**Proposition 5** *The transfers to protection buyers are*

$$\begin{aligned}\tau^B(\bar{\theta}, \bar{s}) &= (E[\tilde{\theta}|\bar{s}] - \bar{\theta}) - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]} A [\alpha(\underline{s})(R - C) + (1 - \alpha(\underline{s}))\mathcal{P}] < 0, \\ \tau^B(\underline{\theta}, \bar{s}) &= (E[\tilde{\theta}|\bar{s}] - \underline{\theta}) - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]} A [\alpha(\underline{s})(R - C) + (1 - \alpha(\underline{s}))\mathcal{P}] > 0,\end{aligned}\tag{17}$$

so that (14) implies the limited liability constraint does not bind in state  $(\underline{\theta}, \bar{s})$ . Furthermore, if the limited liability constraint is slack in state  $(\underline{\theta}, \underline{s})$ , the transfers to protection buyers after a bad signal are

$$\begin{aligned}\tau^B(\bar{\theta}, \underline{s}) &= (E[\tilde{\theta}|\underline{s}] - \bar{\theta}) + A [\alpha(\underline{s}) + (1 - \alpha(\underline{s}))\mathcal{P}] < 0 \\ \tau^B(\underline{\theta}, \underline{s}) &= (E[\tilde{\theta}|\underline{s}] - \underline{\theta}) + A [\alpha(\underline{s}) + (1 - \alpha(\underline{s}))\mathcal{P}] > 0.\end{aligned}\tag{18}$$

Otherwise, the transfers after a bad signal are

$$\begin{aligned}\tau^B(\bar{\theta}, \underline{s}) &= \alpha(\underline{s})A - (1 - \alpha(\underline{s}))A \frac{(1 - \pi)R - \mathcal{P}}{\pi} \\ \tau^B(\underline{\theta}, \underline{s}) &= \alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR > 0.\end{aligned}\tag{19}$$

In the optimal contract, if the limited liability constraint is slack in state  $(\underline{\theta}, \underline{s})$ , then there is full risk sharing given the signal. That is, for a given signal  $s$ , the consumption of the protection buyer is the same irrespective of whether  $\bar{\theta}$  or  $\underline{\theta}$  realizes. On the other hand, in contrast with the first best, transfers vary with the signal. This is because, after a bad signal, it is difficult to provide incentives to the agent. Thus, incentive compatibility reduces the transfers that can be requested from the protection seller. Correspondingly, due to incentive problems, the protection buyer is exposed to signal risk, as her consumption is larger after a good signal than after a bad signal. Cross-subsidization across signals mitigates that effect, but only imperfectly, due to incentive constraints. Cross-subsidization across realizations of the signal is possible because the parties commit to the contract at time 0, before advance information is observed. If the contract was written after that information had been observed, such cross-subsidization would not be possible. This would reduce the scope for insurance, in the line with the Hirshleifer (1971) effect.

To further analyze these effects consider the structure of the transfers in Proposition (5). Each of the transfers in (17) has two components. The first one is the transfer implementing

full risk-sharing conditional on a good signal. The second one reflects cross-subsidization across signals. Transfers in (18) have the same structure except that the first component now reflects full risk-sharing conditional on a bad signal.

The expectation of the first component of these transfers, taken over signals and final realizations of  $\theta$  is 0. This is what would arise with actuarially fair insurance. But the insurance offered by the protection seller is not actuarially fair. It involves a premium, to compensate the protection seller for the efficiency loss induced by margins:  $prob[\underline{s}]\alpha(\underline{s})(R - C - 1)$ . This premium is equal to the expectation of the second component of the transfers in (17) and (18).

The structure of the transfers in (19) is different. When limited liability binds in state  $(\underline{\theta}, \underline{s})$ , full risk-sharing conditional on the signal is no longer possible, as protection sellers' resources in state  $(\underline{\theta}, \underline{s})$  are insufficient. Conditional on a bad signal, the transfers in (19) implement whatever risk-sharing is still possible given the binding limited liability constraint.

Now, turn to the determination of the optimal margin call after a bad signal. We first note that putting all the assets of the protection seller in the margin account cannot be optimal.

**Proposition 6**

$$\alpha^*(\underline{s}) < 1. \tag{20}$$

The logic underlying Proposition 6 is the following. When assets are put in the margin account, they earn lower return than under the management of the protection seller exerting effort. This reduces the resources available to pay insurance to the protection buyer. To cope with this dearth of resources, when  $\alpha^*(\underline{s}) = 1$  all the assets in the margin account must be transferred to the protection buyer when  $\underline{\theta}$  realizes. In this case, as can be seen by inspecting (19) for  $\alpha^*(\underline{s}) = 1$ , the structure of transfers is highly constrained. In fact, it is so constrained that very little risk sharing can be achieved. Hence, a contract requesting  $\alpha^*(\underline{s}) = 1$  is suboptimal.

To analyze the precise amount of margin the calls, it is useful to consider the ratio of the marginal utility of a protection buyer after a bad and a good signal. Denoting this ratio by

$\varphi$ , we have

$$\varphi = \frac{u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))} \quad (21)$$

In the first-best, there is full insurance and  $\varphi$  is equal to 1. With moral hazard, the protection buyer is exposed to signal risk. This makes insurance imperfect and drives  $\varphi$  above one.

Given the transfers in Proposition 5,  $\varphi$  is a known function of exogenous variables and  $\alpha(\underline{s})$ . (17) implies that  $\tau^B(\bar{\theta}, \bar{s})$  is decreasing in  $\alpha(\underline{s})$ . Hence the denominator of  $\phi$  is increasing in  $\alpha(\underline{s})$ . On the other hand, the numerator of is decreasing in  $\alpha(\underline{s})$  (irrespective of whether the limited liability condition in state  $(\underline{\theta}, \bar{s})$  binds or not). Hence,  $\varphi$  is decreasing in  $\alpha(\underline{s})$ . Higher margins reduce  $\varphi$ , as they reduce the wedge between consumption after a good signal and after a bad one, i.e., they improve insurance against signal risk. Optimal margins tradeoff this benefit with their cost: assets in the margin account are less profitable than under the management of the protection seller exerting effort. This tradeoff gives rise to the following proposition.

**Proposition 7** *If  $\mathcal{P} > 1$ , margins are not used. Otherwise, we have the following: If  $\varphi(0) < 1 + \frac{R-C-1}{1-\mathcal{P}}$ , then it is optimal not to use margins. Otherwise, there are two cases. If*

$$\varphi\left(1 - \frac{\pi\Delta\theta}{A(R-\mathcal{P})}\right) < 1 + \frac{R-C-1}{1-\mathcal{P}}, \quad (22)$$

*the limited liability constraint is slack in state  $(\underline{\theta}, \underline{s})$  and the optimal margin solves*

$$\varphi(\alpha^*(\underline{s})) = 1 + \frac{R-C-1}{1-\mathcal{P}}, \quad (23)$$

*while, if (22) does not hold, the optimal margin solves*

$$\varphi(\alpha^*(\underline{s})) = 1 + \frac{R-C-1}{1-\mathcal{P}} + \frac{1-\pi}{1-\mathcal{P}} \frac{u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) - u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))}. \quad (24)$$

The right-hand side of (23) reflects the tradeoff between the costs and benefits of margins. The numerator,  $R - C - 1$ , is the opportunity cost of depositing a margin. The denominator goes up as  $\mathcal{P}$  decreases, i.e., as the incentive problem gets more severe.

When margins are as in (23), consistency requires that there be enough resources to provide full insurance conditional on the signal. This is the case if (22) holds. Consistent

with intuition, this is the case if  $R$  is large enough. When there is full risk sharing conditional on the signal, the last term on the right hand-side of (24) is 0. In that case, (24) simplifies to (23). This case is illustrated in Figure 3. The figure is useful to examine graphically the effect an increase in  $p$ , reducing pledgeable income  $\mathcal{P}$ . The decrease in  $\mathcal{P}$  shifts curve  $\varphi$  upwards while shifting  $1 + \frac{R-C-1}{1-\mathcal{P}}$  downwards. This raises the optimal margin in (23). When incentive problems become more severe, margins are needed more, to relax the incentive constraint.

**Insert Figure 3 here**

On the other hand, when the limited liability constraint binds in state  $(\underline{\theta}, \underline{s})$ , full risk-sharing conditional on the signal is not achievable, so that  $u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) > u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))$ . The last term on the right hand-side of (24) is strictly positive, and, correspondingly, margins are lower than when the limited liability condition is slack. Again, this is because (taking as given that there is effort) margins reduce the amount of resources eventually available to pay insurance. When limited liability binds, these resources are sorely needed. So it is preferable to reduce margins, in order to increase the amount of resources available. The following corollary gives a sufficient condition for (22) to hold.

**Corollary 2** *A sufficient condition for the limited liability condition to be slack in state  $(\underline{\theta}, \underline{s})$  is*

$$1 - \frac{\pi\Delta\theta}{A(R - \mathcal{P})} > \frac{(1 - \pi)R - \mathcal{P}}{\underline{\pi} + (1 - \pi)R - \mathcal{P}}. \quad (25)$$

Condition (25) holds if  $\underline{\pi}\Delta\theta$  is not too large. In that case, full risk-sharing after a bad signal does not request too large resources, and can thus be implemented.

In the first-best the transfers depend only on the realization of  $\theta$  and the optimal contract can be implemented with a simple forward contract. In contrast, with moral-hazard and risk-management effort after both signals, the transfers depend on the realizations of  $\theta$  and  $s$ . The optimal contract can be implemented by the sale of a forward contract on the underlying asset  $\theta$  by protection buyers (as in the first-best) together with the purchase of a forward contract on the signal  $s$ . The forward contract on  $s$  generates a gain for protection sellers

in state  $\underline{s}$ . This gain increases their pledgeable income after a bad signal and thus restores incentive compatibility in the light of the liability from the forward contract on  $\theta$ .<sup>12</sup>

One may wonder whether the optimal contract that is contingent on the signal  $s$  can be replicated by renegotiating - after  $s$  is observed - a contract that is initially independent of the signal,  $\tau^B(\theta)$ . Suppose, for example, that the parties initially agree on the transfers  $\tau^B(\theta) = \tau^B(\theta, \bar{s})$ . These are the optimal transfers in case of a good signal so there is no scope for renegotiation after observing  $\bar{s}$ . But what about after observing a bad signal  $\underline{s}$ ? Is it a Pareto improvement to renegotiate and switch to the optimal contract after a bad signal,  $\tau^B(\theta, \underline{s})$ ?

Sticking to  $\tau^B(\theta, \bar{s})$  after a bad signal violates protection sellers' incentive compatibility constraint. They do not exert risk-management effort, fail with probability  $1 - p$  and expect to obtain

$$\underline{\pi}p (AR - \tau^B(\bar{\theta}, \bar{s})) + (1 - \underline{\pi})p (AR - \tau^B(\underline{\theta}, \bar{s})) .^{13}$$

Although the expected payoffs from  $\tau^B(\theta, \bar{s})$  render the contract an asset when the good signal occurred, the expected payoffs from this contract after a *bad* signal render the contract a liability. The only benefit of sticking with the contract is that it avoids payment to protection buyers with probability  $1 - p$ .

If protection sellers switch to  $\tau^B(\theta, \underline{s})$ , they exert risk-management effort and expect to obtain

$$\underline{\pi} (AR - \tau^B(\bar{\theta}, \underline{s})) + (1 - \underline{\pi}) (AR - \tau^B(\underline{\theta}, \underline{s})) - AC$$

By switching, protection sellers increase the expected payoff on their assets since effort is more productive than no effort (see condition (1)). Moreover, switching reduces the payment to the protection buyers as  $\tau^B(\underline{\theta}, \underline{s}) < \tau^B(\underline{\theta}, \bar{s})$ .

Substituting for the transfers and re-arranging, it follows that protection sellers renegotiate if and only if

$$AP < E[\theta] - \text{prob}[\bar{s}]E[\tilde{\theta}|\bar{s}] \tag{26}$$

which is always satisfied since  $AP < E[\theta] - E[\tilde{\theta}|\underline{s}]$  (see Lemma 1).

<sup>12</sup>While this implementation is plausible, it is not unique. Other financial contracts with gains for protection sellers after  $\underline{s}$  such as options can be used.

<sup>13</sup>Recall that with probability  $1 - p$ ,  $\tilde{R}_j = 0$  and  $\tau^B(\theta, \tilde{s}, 0) = 0$ .

For protection buyers, the renegotiation decision is determined by two factors. First, sticking to  $\tau^B(\theta, \bar{s})$  after a bad signal implies higher transfers from the CCP. But - and this is the second factor - it exposes them to counterparty risk. With the CCP insuring against counterparty risk, there is no downside to sticking to the original contract as the protection buyers do not internalize the benefits of the switch to  $\tau^B(\theta, \underline{s})$ . Thus, with a CCP, the optimal contract cannot be implemented by renegotiation. But suppose trading occurs over-the-counter and contracting is bilateral. Then a protection buyer is exposed to the potential failure of his protection seller when he does not renegotiate the contract after a bad signal and his expected utility is

$$\pi u(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s})) + (1 - \pi) p u(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}))$$

If the protection buyer renegotiates, his expected utility is

$$\pi u(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s})) + (1 - \pi) u(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})).$$

Substituting for the transfers, a protection buyer renegotiates the bilateral contract when counterparty risk is sufficiently large:

$$p < \frac{\frac{u(E(\theta|\underline{s}) + AP)}{u(E(\theta|\bar{s}) - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]} AP)} - \pi}{1 - \pi}$$

If the condition holds, then the optimal bilateral contract can be implemented by a simple renegotiation game in which the contract itself does not depend on the realization of the signal. At time 0, each protection buyer makes a protection seller a take-it-or-leave-it offer and at time 1, after observing the realization of the signal, he can make another take-it-or-leave-it offer. In this game, a protection buyer finds it optimal to offer  $\tau^B(\theta, \bar{s})$  at time 0 and  $\tau^B(\theta, \underline{s})$  at time 1 if  $\underline{s}$  is observed.

### 3.2 No effort after a bad signal (risk-taking)

Incentive compatibility after a bad signal reduces risk-sharing. Protection buyers may find this reduction in insurance too costly. They may instead choose to accept shirking on risk prevention effort (risk-taking) by protection sellers in exchange for a better sharing of the risk

associated with  $\tilde{\theta}$ . In this subsection, we characterize the optimal contract with risk-taking after a bad signal.

After a good signal, protection sellers exert risk-management effort so that  $\tilde{R}_j = R$  for all  $j$ . After a bad signal, protection sellers do not exert risk-management effort so that  $\tilde{R}_j = R$  for a proportion  $p$  of sellers and  $\tilde{R}_j = 0$  for a proportion  $1 - p$  of sellers. Hence, the transfer  $\tau^S$  from the CCP to a protection seller must now be contingent on the realization of  $\tilde{R}_j$ . By contrast, the transfer  $\tau^B$  from the CCP to a protection buyer does not have to be contingent on the realization of a particular  $\tilde{R}_j$ . The CCP can mutualize counterparty risk and provide insurance to risk-averse protection buyers. However, the aggregate amount of resources protection sellers generate differs after a good signal and after a bad signal. After a bad signal, only proportion  $p$  of protection sellers generate return  $R$  while proportion  $1 - p$  of sellers generate a zero return and cannot make any payments to the CCP as they are protected by limited liability.

The CCP chooses transfers to buyers and sellers,  $\tau^B(\tilde{\theta}, \tilde{s}, \tilde{R})$  and  $\tau^S(\tilde{\theta}, \tilde{s}, \tilde{R})$ , to maximize buyers' utility

$$\begin{aligned} & \pi \lambda u(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}, R)) + (1 - \pi)(1 - \lambda)u(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}, R)) \\ & + \pi(1 - \lambda)u(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, pR)) + (1 - \pi)\lambda u(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s}, pR)) \end{aligned} \quad (27)$$

where, after a bad signal,  $\tau^B$  is written as a function of  $pR$  to indicate mutualization of counterparty risk by the CCP.

The feasibility constraints of the CCP after good and a bad signal, respectively, are given by

$$\tau^B(\theta, \bar{s}) \leq -\tau^S(\theta, \bar{s}, R) \quad \forall(\theta, \bar{s}) \quad (28)$$

$$\tau^B(\theta, \underline{s}) \leq -p\tau^S(\theta, \underline{s}, R) - (1 - p)\tau^S(\theta, \underline{s}, 0) \quad \forall(\theta, \underline{s}) \quad (29)$$

The limited liability constraints for sellers whose assets generate  $R$  and for those whose assets generate 0, respectively, are given by:

$$-\tau^S(\theta, s, R) \leq \alpha(s)A + (1 - \alpha(s))AR \quad \text{for } R_j = R \quad (30)$$

$$-\tau^S(\theta, s, 0) \leq \alpha(s)A \quad \text{for } R_j = 0 \quad (31)$$



The seller's incentive constraint after a good signal is, as before,

$$\alpha(\bar{s})A + (1 - \alpha(\bar{s}))A\mathcal{P} \geq -E[\tau^S(\theta, \bar{s}, R)], \quad (32)$$

whereas after a bad signal, the seller must prefer not to exert effort

$$\begin{aligned} E[\tau^S(\theta, \underline{s}, R)] + \alpha(\underline{s})A + (1 - \alpha(\underline{s}))A(R - C) \leq \\ pE[\tau^S(\theta, \underline{s}, R)] + (1 - p)E[\tau^S(\theta, \underline{s}, 0)] + \alpha(\underline{s})A + (1 - \alpha(\underline{s}))pAR, \end{aligned}$$

or, equivalently,

$$(1 - \alpha(\underline{s}))A\mathcal{P} \leq -E[\tau^S(\theta, \underline{s}, R)] + E[\tau^S(\theta, \underline{s}, 0)]. \quad (33)$$

Finally, the seller's participation constraint with risk-taking is

$$\begin{aligned} \text{prob}[\bar{s}]\alpha(\bar{s})A(R - C - 1) + \text{prob}[\underline{s}]\alpha(\underline{s})A(pR - 1) + \text{prob}[\underline{s}](1 - p)A\mathcal{P} \\ \leq \text{prob}[\bar{s}]E[\tau^S(\theta, \bar{s}, R)] + \text{prob}[\underline{s}](pE[\tau^S(\theta, \underline{s}, R)] + (1 - p)E[\tau^S(\theta, \underline{s}, 0)]) \end{aligned} \quad (34)$$

The expected transfer from the CCP to a protection seller (right-hand side) is positive. If a seller enters the position, she must be compensated for the potential efficiency loss (left-hand side). The loss is due to two factors: 1) costly margins after good and a bad signal (where  $R - C - 1$  is the opportunity cost of margins when a seller exerts effort and  $pR - 1$  is the opportunity cost of margins when she does not) and 2) the loss of pledgeable income in the event of default, which occurs with probability  $\text{prob}[\underline{s}](1 - p)$ . Thus, the contract with no effort after a bad signal is actuarially unfair. The higher the pledgeable income, the greater the efficiency loss generated by risk-taking after a bad signal, the more actuarially unfair the contract.

We can re-write the seller's participation constraint with risk-taking as

$$\begin{aligned} A\text{prob}[\bar{s}]\alpha(\bar{s})(R - C - 1) + A\text{prob}[\underline{s}][R - C - (\alpha(\underline{s}) + (1 - \alpha(\underline{s}))pR)] \\ \leq \text{prob}[\bar{s}]E[\tau^S(\theta, \bar{s}, R)] + \text{prob}[\underline{s}](pE[\tau^S(\theta, \underline{s}, R)] + (1 - p)E[\tau^S(\theta, \underline{s}, 0)]) \end{aligned} \quad (35)$$

On the left-hand side, there is again the efficiency loss from entering the contract with risk-taking. After a good signal, the seller exerts effort but there is an opportunity cost of margins, given by  $R - C - 1$ . After a bad signal, the seller does not exert effort and

the efficiency loss is given by the difference between  $R - C$ , the return on assets when not entering the contract and doing effort, and  $\alpha(\underline{s}) + (1 - \alpha(\underline{s}))pR$ , the expected return under the contract with risk-taking.

We first show that in the optimal contract with risk-taking, the feasibility constraints and the participation constraint must bind, i.e., protection sellers earn no rents and all resources available for insurance are passed on to protection buyers.

**Proposition 8** *In the optimal contract with risk-taking after a bad signal, the feasibility constraints bind for all  $(\theta, s)$  and the participation constraint binds.*

The next proposition characterizes the use of margins in the contract with risk-taking and narrows down the parameter space for which risk-taking after a bad signal can be optimal:

**Proposition 9** *In the optimal contract with risk-taking after a bad signal, margins are not used after signal  $\bar{s}$  if the incentive constraint given  $\bar{s}$  is slack or if the moral-hazard is not severe, i.e.,  $\mathcal{P} \geq 1$ . After signal  $\underline{s}$ , margins are not used if  $pR \geq 1$ . If  $pR < 1$ , then  $\alpha^*(\underline{s}) = 1$ . Such contract is, however, dominated by the one with effort after a bad signal.*

Without effort after a bad signal, the expected per-unit return on the seller's balance sheet is  $pR$ . If  $pR < 1$ , this is lower than what assets return on the margin account. Hence, it is more profitable to deposit all of the protection seller's assets in the margin account,  $\alpha = 1$ , where they earn a greater return and are ring-fenced from moral hazard. But protection buyers can do at least as well by requesting effort after a bad signal since, there too,  $\alpha = 1$  can be selected (but, as we know from Proposition 6, it is never optimal). It follows that the contract with margins and no effort after a bad signal can only be strictly optimal if  $pR \geq 1$ .

The next proposition characterizes the optimal transfers in the contract with risk-taking after a bad signal.

**Proposition 10** *If  $pR < 1$ , then risk-taking is suboptimal. Otherwise, the optimal contract with risk-taking after a bad signal provides full insurance to protection buyers if and only if*

$$pAR \geq \pi\Delta\theta - (1 - p) \text{prob}[\underline{s}]AP. \quad (36)$$

The transfers are given by

$$\begin{aligned}\tau^B(\bar{\theta}, \bar{s}) &= \tau^B(\bar{\theta}, \underline{s}) = -(1 - \pi)\Delta\theta - \text{prob}[\underline{s}](1 - p)A\mathcal{P} < 0, \\ \tau^B(\underline{\theta}, \bar{s}) &= \tau^B(\underline{\theta}, \underline{s}) = \pi\Delta\theta - \text{prob}[\underline{s}](1 - p)A\mathcal{P} > 0.\end{aligned}$$

In contrast to the contract with effort after a bad signal, the contract with risk-taking does not react to the signal, i.e.,  $\tau^B(\bar{\theta}, \bar{s}) = \tau^B(\bar{\theta}, \underline{s})$ . The consumption of the buyer is equalized across states (i.e., there is full insurance, as in the first-best) as long as the amount of resources generated under risk-taking (by the protection sellers who succeed), equal to  $pAR$ , is sufficiently high. However, since protection buyers must compensate protection sellers for the efficiency loss due to risk-taking (given by the loss of pledgeable income in the event of default after a bad signal,  $\text{prob}[\underline{s}](1 - p)A\mathcal{P}$ ), the consumption of protection buyers falls short of the first-best consumption levels. Condition (36) ensures that the limited liability constraints are slack under full insurance. On the left-hand side are the aggregate resources generated by protection sellers. On the right-hand side is the transfer that would be paid in the first-best, minus the payment requested by protection-sellers to offset the efficiency loss they incur due to risk-taking.

Risk-taking can be optimal only if it is not too inefficient, i.e., if  $pR \geq 1$ . In that case, margins are not used. Since protection sellers engage in risk-taking after a bad signal, margins do not help with incentives. Margins are also not needed to insure buyers against counterparty risk since it is mutualized by the CCP. Thus, mutualization tackles ex-post counterparty risk in the contract with risk-taking, while margins tackle ex-ante incentives in the contract with effort.

Condition (36) can be re-written as

$$A[(1 - \text{prob}[\underline{s}])pR + \text{prob}[\underline{s}](R - C)] \geq \pi\Delta\theta.$$

Since  $R - C > 1$  and  $pR \geq 1$  (the latter condition is necessary for the contract with risk-taking to be optimal), it follows that

$$A[(1 - \text{prob}[\underline{s}])pR + \text{prob}[\underline{s}](R - C)] > A.$$

Hence, a sufficient condition for (36) to hold is

$$A \geq \pi\Delta\theta. \tag{37}$$

In the optimal contract with risk-taking after bad news, thanks to the mutualization of counterparty risk by the CCP, transfers are not contingent on signals or on individual protection seller's returns. Hence the optimal contract can be implemented with a single forward contract (as in the first-best) provided it is insured by the CCP (unlike in the first-best). The forward contract, however, is sold at a discount relative to the expected value of the underlying risk, in order to compensate the protection sellers for the loss of pledgeable income in default.

### 3.3 Risk-sharing and risk-taking

The contract under which protection sellers exert effort after both signals entails limited risk-sharing for buyers but entails no risk-taking by sellers (Subsection 3.1), while the contract with no effort after a bad signal entails full risk-sharing for protection buyers but is actuarially unfair and falls short of the first-best due to the loss of resources in default (Subsection 3.2). The next proposition characterizes the optimal choice between the two contracts as a function of the probability of success under risk-taking,  $p$ .

**Proposition 11** *Assume (37) holds. There exists a threshold value of the success probability under no effort  $\hat{p}$  such that risk-prevention effort after bad news is optimal if and only if  $p \leq \hat{p}$ .*

The logic of the proposition is illustrated in Figure 4. Consider the expected utility of the protection buyer when effort is requested after bad news. It decreases when  $p$  increases. For this contract, indeed, the only effect of an increase in  $p$  is to tighten the incentive constraint, and thus reduce risk-sharing. Now turn to the expected utility of the protection buyer when effort is not requested after bad news. In contrast with the previous case, it increases when  $p$  increases. Indeed, for this contract, the only effect of an increase in  $p$  is to increase the amount of resources available after bad news. Hence the result, stated in the proposition, that risk-prevention effort after bad news is optimal if and only if  $p$  is lower than a threshold.

**Insert Figure 4 here**

## 4 Non-linear cost of risk-management effort

In this Section, we examine a non-linear cost of risk-management effort. Specifically, we assume a per-unit cost of

$$C = c + \gamma(1 - \alpha(s))A, \quad (38)$$

where  $(1 - \alpha(s))A$  denotes assets under management after a margin call at  $t = 1$ . Previously, we had  $\gamma = 0$ .

With  $\gamma = 0$  and  $\mathcal{P} < 1$ , margins relax the incentive constraint and allow more risk-sharing. We show that this result still holds, and actually strengthens, when the cost of risk-management effort is convex, i.e. when  $\gamma > 0$ .<sup>14</sup>

We showed that larger margins ring-fence assets from moral-hazard in risk-management. If the cost is convex ( $\gamma > 0$ ), there is an additional effect: the cost of controlling the risk of those assets still under the management of protection seller decreases when margins increase. If cost is concave ( $\gamma < 0$ ), however, then this cost *increases*, which works against the ring-fencing effect. Moreover, the optimization problem may become ill-behaved when the cost of risk-management effort is concave.

Since margins do not play any role in the contract without risk-management effort after a bad signal, we only need to re-consider the contract with effort. We know from our previous analysis (Section 3.1) that the feasibility constraints as well as the participation constraint bind: there is no reason to have idle resources or to leave rents to protection sellers.<sup>15</sup> Moreover, the incentive constraint is slack after a good signal (and there is no margin call) while it binds after a bad signal, in which case there may be a margin call. This logic does not depend on the shape of the cost of effort function.

The (binding) incentive constraint after a bad signal now is:

$$\begin{aligned} E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] + \alpha(\underline{s})A + (1 - \alpha(\underline{s}))A(R - c - \gamma(1 - \alpha(\underline{s}))A) \\ = p \left( E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] + \alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR \right), \end{aligned}$$

---

<sup>14</sup>Berk and Green (2004) model the active management of a portfolio of financial assets and assume the cost of doing so is increasing and convex in the amount of funds under management (which corresponds to  $\gamma > 0$ ). For example, as the size of an active fund grows it becomes more and more difficult to gather the information to add value over a passively management fund.

<sup>15</sup>For brevity, we ignore the limited liability constraints in this extension. This does not affect our conclusions.

which, as in Section 3, simplifies to

$$\alpha(\underline{s})A + (1 - \alpha(\underline{s}))A\mathcal{P}(\alpha(\underline{s})) = E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}], \quad (39)$$

where the pledgeable return now depends on the size of the margin call after a bad signal:

$$\mathcal{P}(\alpha(\underline{s})) \equiv R - \frac{c + \gamma(1 - \alpha(\underline{s}))A}{1 - p}. \quad (40)$$

Margins improve risk-sharing when they relax the incentive constraint after a bad signal (39). With a non-linear cost of effort, this is the case when

$$\mathcal{P}(\alpha(\underline{s})) - (1 - \alpha(\underline{s}))\mathcal{P}' < 1. \quad (41)$$

In the case of a linear cost of effort,  $\gamma = 0$ , we have  $\mathcal{P}(\alpha(\underline{s})) = R - \frac{c}{1-p}$  and  $\mathcal{P}' = 0$  so that condition (41) simplifies to the condition on the severity of the moral-hazard  $\mathcal{P} < 1$  in Section 3.1. When the cost of effort is convex,  $\gamma > 0$ , it is easier to satisfy condition (41). Margins now have an additional benefit. They reduce the cost of managing the risk of those assets still under the control of protection sellers. The opposite happens when the cost of effort is concave,  $\gamma < 0$ . Then it is more difficult to satisfy (41).

To determine the optimal margin when the cost of effort is non-linear, we proceed as in Section 3.1. We focus on the convex cost case since we cannot be sure the first-order condition identifies a local maximum when the cost is concave (see the proof of Proposition 12 for details). The transfers  $\tau^B$  and  $\tau^S$  have the same structure as in Proposition 5, except that  $\mathcal{P}$  is now given by (40). We obtain the following Proposition (where, as before,  $\varphi$  denotes the ratio of the marginal utility of a protection buyer after a bad and a good signal).

**Proposition 12** *With a convex cost of risk-management effort,  $\gamma > 0$ , an optimal interior margin after a bad signal  $\alpha^*(\underline{s})$  is given by*

$$\varphi(\alpha^*(\underline{s})) = 1 + \frac{R - C - 1}{1 - [\mathcal{P}(\alpha^*(\underline{s})) - (1 - \alpha^*(\underline{s}))\mathcal{P}']}.$$

As in Proposition 7, the optimal interior margin solves the tradeoff between more risk-sharing across signals and the opportunity cost of margin deposits. Unlike in Proposition 7, however, the extra pledgeable return of margins (relative to those assets still under the

control of protection sellers) now depends on the size of the margin call. With a convex risk-management cost, larger margins make it easier to manage the risk of those assets still under the control of protection sellers. This makes margins more attractive compared to the linear cost of effort case. Consequently, we obtain the following comparative static result:

**Proposition 13** *Suppose  $\gamma \geq 1$  and  $1 > R - \frac{c}{1-p}$ . The stronger decreasing returns to scale in risk-management lead to larger margin calls after a bad signal,  $\frac{d\alpha^*(s)}{d\gamma} > 0$ .*

With a convex cost of risk-management effort, margins not only ring-fence assets from moral-hazard but also have a beneficial scale effect for the risk management of those assets still subject to moral-hazard. Our baseline model did not have any scale effects and focused exclusively on the ring-fencing effect of margins.

## 5 Empirical implications

According to our theory, a strong and pledgeable asset base ( $AP$ ) maintains risk-management incentives in the presence of derivative exposures.<sup>16</sup> Asset pledgeability decreases when the cost of risk-management effort increases. Pledgeability therefore suffers when insufficient resources are devoted to risk-management or when the power of risk-managers is limited. Pledgeability also suffers when financial institutions and their activities become more complex and opaque. Consistent with these predictions, Ellul and Yerramili (2013) find cross-sectional differences in banks' risk-taking behaviour as a function of the strength and independence of banks' risk-management. Specifically, U.S. bank holding companies with a higher Risk Management Index (indicating better risk-management) before the onset of the 2007-09 financial crisis have lower tail-risk exposures and better overall performance during the financial crisis.

Asset pledgeability also depends on the macroeconomic and financial environment in which financial institutions operate. For example, an environment characterized by a low probability of failure even when there is no risk-management effort (high  $p$ ) can be viewed as a "benign"/low-risk economic situation. In our analysis, derivatives contracts that offer

---

<sup>16</sup>While for simplicity protection sellers have no initial debt in our model, to gauge this implication empirically one should consider assets net of liabilities.

ample insurance but undermine risk-management incentives are (privately) optimal in such a benign environment (see Proposition 11). This implication resonates with the idea of excessive risk building up in “good” times - an idea that figures prominently in accounts of the recent global financial crisis (see for example Borio, 2011). [note that the build up of leverage in good times also features prominently in Acharya-Vish, except that the mechanism is different: something to mention in the intro?]

A related implication of our model is contagion across seemingly unrelated assets when bad shocks hit (see, e.g., Billio et al., 2012). The reason why there is contagion in our analysis is the desire to hedge risk. Protection buyers want to diversify the aggregate component of their risk and thus turn to protection sellers. Conditional on protection sellers’ risk-management effort, their balance-sheet risk is independent, which makes them suitable insurers of protection buyers’ risk. Moreover, in a benign economic environment protection buyers seek ample insurance and accept the lack of risk-management incentives of protection sellers in case of negative news about the insured risk. When such news do occur, e.g., slowing house prices when mortgage default is insured, then protection sellers become more likely to default. One would then observe an increased correlation between the mortgage values and the values of financial institutions without direct exposure to mortgage default.

Another set of factors affecting risk-management incentives relates to the type of derivatives contracts sold by protection sellers. Since a protection seller’s incentives are jeopardized when a derivative position turns into an expected liability for her, the accuracy of advance information about how the underlying asset will perform matters. In particular, an increase in the precision of the advance signal ( $\lambda$ ) unambiguously worsens the incentive problem of protection sellers. For example, information about the performance of mortgage-backed securities and CDS contracts written on them was unavailable before 2006 even though the issuance of mortgage-backed securities was around \$ 2 trillion in every year from 2002 until 2006 (see, e.g., Fender and Scheicher, 2008). The ABX.HE indices that provide this information were introduced only in January 2006. Moreover, as of early 2007, the prices for the index on AAA securitizations and those on BBB securitizations, which were virtually identical until then, started to diverge. In the light of our analysis, this increase in the precision of information made it more difficult to have incentive-compatible insurance of mortgage



defaults and contributed to the development of the financial crisis: To the extent that ample insurance kept being written, it came at the expense of risk-taking incentives for insurers.

In addition to the quality of advance information about the asset underlying the derivative, the structure of the derivative's payoffs plays a role as well. First, more risk to be hedged (higher  $\Delta\theta$ ) worsens the incentive problem, *ceteris paribus*. For example, an interest-rate swap, where only interest payments are at stake, hedges less risk and leads to less incentive problems than a credit-default swap, where also the principal payment is at stake.

Second, the symmetry of the hedged risk matters, too. For example, while the payoff from an interest rate swap is more like a symmetric coin-flip, the payoff from a credit-default swap is highly skewed: most of the time, protection sellers collect a small insurance premium but in the rare case of default, they have to make large payments to protection buyers. To analyze the effect of an increase in the skewness of the hedged risk on incentives, we increase the probability  $\pi$  of a good outcome for the protection buyer's risk  $\theta$  while keeping its mean and the standard deviation constant.<sup>17</sup> An increase of  $\pi$  increases the amount of risk to be hedged,  $\Delta\theta$ . Consequently, protection buyers demand more insurance, which increases the incentive problem for protection sellers. There is, however, a countervailing effect when the skewness  $\pi$  is already large. In that case, the good outcome of the hedged risk is quite likely and the information content of a bad signal  $\underline{s}$  is low. At high levels of  $\pi$ , a further increase of skewness mutes the negative effect of bad news on incentives.<sup>18</sup> As long as  $\pi < \lambda$  (the precision of the signal  $\underline{s}$ ), the negative effect on incentives from larger amounts of risk dominates and more skewness leads to more severe incentive problems. In sum, it is more difficult to maintain risk-management incentives when protection sellers insure credit risk compared to interest rate risk.

Finally, the use of margins depends on their opportunity cost and the degree to which they alleviate protection sellers' incentive problem. The opportunity cost of margins depends on the risk-free rate (normalized to one in our analysis) since this is the rate assets on the

---

<sup>17</sup>The mean  $\mu$  and the standard deviation  $\sigma$  of  $\tilde{\theta}$  are  $\mu = \pi\bar{\theta} + (1 - \pi)\underline{\theta}$  and  $\sigma = \sqrt{\pi(1 - \pi)}\Delta\theta$ , respectively. We can therefore write  $\bar{\theta}$  and  $\underline{\theta}$  as a function of  $\pi$  as follows:  $\bar{\theta} = \mu + \sigma\sqrt{\frac{1 - \pi}{\pi}}$  and  $\underline{\theta} = \mu - \sigma\sqrt{\frac{\pi}{1 - \pi}}$ . Holding the mean and standard deviation constant, an increase in  $\pi$  leads to more skewness (when  $\pi > \frac{1}{2}$ ).

<sup>18</sup>Formally, the size of the incentive problem depends on  $E[\theta] - E[\theta|\underline{s}]$ , which is the product of  $\Delta\theta$  and  $(\pi - \underline{\pi})$ , where the latter captures the effect of bad news.

margin account earn. When risk-free rates are low compared to the return on productive investment opportunities, the opportunity cost of margins increases and the optimal margin is lower. In terms of alleviating the incentive problem, margins are particularly beneficial when the cost of risk-management effort is convex, and the optimal margin is higher the more convex risk-management costs are (see Proposition 13).

## 6 Policy implications

### 6.1 Margins and equity capital

We showed that margins allow for more incentive-compatible insurance as they ring-fence assets from protection seller moral-hazard. What could be alternative mechanisms to reduce moral-hazard? In particular, what about a capital requirement so that protection sellers have “skin-in-the-game”? What are the similarities and the differences between margins and equity capital in the context of our analysis? These questions are particularly relevant since the regulatory overhaul in the aftermath of the 2007-2009 financial crisis includes both margin and capital requirements.

Our optimal contracting framework in fact allows for equity capital. At  $t = 0$ , protection sellers have assets  $A$  and no liabilities. Hence, their equity capital (the difference between assets and liabilities) is equal to  $A$ . At  $t = 1$ , after a good signal, the derivative position is an expected asset for a protection seller and the value of her equity increases. After a bad signal, however, the derivative position is an expected liability for a protection seller. The optimal contract with effort limits this liability to  $A[\alpha(\underline{s}) + (1 - \alpha(\underline{s}))\mathcal{P}]$  to preserve protection seller’s incentives to exert risk-management effort (see (15)). The value of a protection seller’s equity capital after a bad signal at  $t = 1$  is

$$(1 - \alpha(\underline{s})) (R - \mathcal{P}) A > 0, \tag{42}$$

which is the difference between the value of protection seller’s assets,  $A[\alpha(\underline{s}) + (1 - \alpha(\underline{s}))R]$ , and the value of her liability. Therefore, one interpretation of the optimal contract with effort is that it does require protection sellers to hold a minimum amount of equity.

If margin calls could not be made (e.g., if there was no enforcement mechanism for margins), then the amount of equity protection sellers must hold would increase (see (42)).

Margins allow protection sellers to engage in incentive-compatible financial contracting with less equity. Margins improve incentives by making the asset side of the balance sheet less susceptible to moral-hazard. With less moral-hazard, the assets can support larger liabilities.

Another way to increase the size of the incentive-compatible liability after a bad signal, and thus increase risk-sharing, is to increase the size of a protection seller's balance-sheet  $A$ . Since protection sellers are 100% equity financed, more equity capital upfront increases the amount of incentive-compatible insurance.

The cost of using margins is the net loss of the return  $R - C - 1$  on the assets a protection seller no longer controls. If increasing equity capital was costless for protection sellers, then they would of course substitute capital for margins. But increasing equity capital is costly for a number of reasons. Issuing equity carries well-known dilution costs. New equity-holder may also demand control rights that reduce the value of existing equity-holders' stake. Finding new equity capital can be time-consuming. Finally, there can be substantial transaction costs when issuing new equity.

Both margins and equity capital can relax the incentive constraint of protection sellers. But there are differences. Margins only improve incentives when the moral-hazard problem is sufficiently severe ( $\mathcal{P} < 1$ ). More equity capital (larger  $A$ ) always improves incentives. Moreover, margins, unlike equity, are linked to individual derivative positions. A margin call only occurs when the derivative position turns into a liability (which depends on information about the underlying asset). An increase in equity is not contingent on the development of derivative positions. This can be quite wasteful since the equity capital is present also when derivative positions are assets. Finally, higher equity capital also benefits other existing claim-holders. Margins in contrast are ear-marked to cover the risk of the specific derivative positions.

## 6.2 CCP design

The key benefit of CCPs in our paper is the mutualization of counterparty risk. Trading partners are therefore more willing to accept counterparty risk when trades are centrally cleared than when they occur over-the-counter (OTC). Nevertheless, the optimal contract and margins in our model are the constrained-efficient outcome on which a regulator cannot

improve.

But suppose we modify the model slightly and add the possibility of losses  $L \geq 0$  when protection sellers default. Such losses can occur when protection sellers have other claim-holders (depositors, debt-holders) or when their default spills over to other parts of the financial system (e.g., via disruptions in interbank markets or payment systems). Social (utilitarian) welfare is given by the sum of protection buyers' utility and the protection sellers' profits net of expected losses  $L$ . Since protection sellers are protected by limited liability, the cost of their default  $L$  will not be taken into account in the privately optimal design of the contract and margins. We can therefore think of  $L$  as a reduced-form measure of the externality a default of protection sellers imposes on the rest of the financial system.

Since the losses occur more often when protection sellers fail to exert risk-management effort (risk-taking), there exists a threshold level of losses,  $L^*$ , such that for losses larger than  $L^*$ , social optimality requires the avoidance of risk-taking even though it may be privately optimal. Because of the mutualization of counterparty risk, risk-taking is more attractive and the case for regulation is *stronger* for centrally cleared than for OTC trades.

The mutualization of counterparty risk through a CCP can lead to free-riding among protection buyers. Since protection buyers are no longer exposed to risk of protection seller default, they have no incentive to insist on transfers and margins that lead to protect sellers' risk-management effort and reduce the risk of default. The cost of insisting on the optimal contract (e.g., monitoring derivative positions) would be borne privately, while the benefit of better incentives and hence lower risk would be mutualized.

We also showed that protection buyers are less likely to renegotiate a sub-optimal contract when trades are centrally cleared than when they occur bilaterally in an over-the-counter market (see Section 3.1). The insurance against counterparty risk through a CCP hurts protection buyers incentives to renegotiate a contract that has large insurance payments but leads to risk-taking by protection sellers. The lack of renegotiation is optimal from the point of view of a single protection buyer but harmful from the point of view of all protection sellers since it leads to a loss of overall resources when protection sellers default.

To prevent such free-riding problems, it is the CCP and not the contracting parties who must set and mandate the optimal transfers and margin calls, and make them contingent

on the development of derivative positions (e.g, monitor real estate prices for CDS contracts written on mortgage loans or written on institutions with large exposures to such loans). To be up to this task, CCPs must be well-governed and have the resources to carry out the necessary monitoring.

## 7 Conclusion

We analyze optimal contracts in the context of hedging using derivatives. We show how contracts designed to engineer risk-sharing generate incentives for risk-taking. When the position of the protection seller becomes a liability for her, it undermines her incentives to exert risk prevention effort. The failure to exert such effort may lead to the default of the protection seller. Thus, a bad signal about derivative positions can propagate to other lines of business of financial institutions and, when doing so, create endogenous counterparty risk.

When the seller's moral hazard is moderate, margins enhance the scope for risk-sharing. Our emphasis on the positive consequence of margins contrasts with the result that margins can be destabilizing, as shown by Gromb and Vayanos (2002) and Brunnermeier and Pedersen (2009). The contrast stems from differences in assumptions. Gromb and Vayanos (2002) and Brunnermeier and Pedersen (2009) take margin constraints as given and, for these margins, derive equilibrium prices. Greater margins force intermediaries to sell more after bad shocks, which pushes prices down and can generate spirals. In contrast, we endogenize margins, but take as given the value of assets a protection seller deposits on a margin account. It would be interesting in future research to combine the two approaches and study how endogenous margins could destabilize equilibrium prices. Destabilization could arise if the margin requirement that is privately optimal for a protection buyer and his counterparty had external effects on other investors via equilibrium prices, as in Gromb and Vayanos (2002) and Brunnermeier and Pedersen (2009). This would be in line with the analysis of equilibrium investment and asset pricing with endogenous financial constraints by Acharya and Viswanathan (2011) and Lorenzoni (2008).

IMPORTED FROM INTRO:

Gromb and Vayanos (2002) show how margin constraints lead to a destabilizing behavior

of arbitrageurs that exacerbates price volatility. Brunnermeier and Pedersen (2009) show that margins constraints can generate downward price spirals. Margin calls force asset sales, which depress prices and lower asset values, triggering further margin calls. Both these analyses, however, take the existence of margin constraints as given. To the best of our knowledge, our paper is the first to study the emergence of margin deposits as a feature of an optimal contract.

ALSO IMPORTED FROM INTRO

Because the context we study is different, the inefficiencies we uncover differ from those analyzed in a borrowing and lending context. Acharya and Viswanathan (2011) offer an insightful analysis of the equilibrium price at which borrowers liquidate assets and the corresponding fire-sales negative externality. This is beyond the scope of our paper. Instead, we show that optimal contracting between protection buyers and sellers can generate endogenous counterparty risk and contagion from negative shocks on the assets of the protection buyer to those of the protection seller. These inefficiencies do not arise in Acharya and Viswanathan (2011).

## References

- Acharya, V. and A. Bisin, 2011, "Counterparty Risk Externality: Centralized versus Over-the-counter Markets," NBER working paper No. 17000.
- Acharya, V. and S. Viswanathan, 2011, "Leverage, moral hazard and liquidity," *Journal of Finance*, 66 (1), 99-138.
- Allen, F. and E. Carletti, 2006, "Credit Risk Transfer and Contagion," *Journal of Monetary Economics*, 53 (1), 89-111.
- Berk, J. and R. Green, 2004, "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 112 (6), 1269-1295.
- Biais, B. and C. Casamatta, 1999, "Optimal Leverage and Aggregate Investment," *Journal of Finance*, 54 (4), 1291-1323.
- Billio, M., M. Getmansky, A. Lo and L. Pelizzon, 2012, "Econometric measures of connectedness and systemic risk in the finance and insurance sectors," *Journal of Financial Economics*, 104 (3), 535-559.
- Bolton, P. and M. Oehmke, 2011, "Credit Default Swaps and the Empty Creditor Problem," *Review of Financial Studies*, forthcoming.
- Bolton, P. and M. Oehmke, 2011, "Should Derivatives be Senior?," working paper, Columbia University.
- Bolton, P. and D. Scharfstein, 1990, "A Theory of Predation Based on Agency Problems in Financial Contracting," *American Economic Review*, 80, 93-106.
- Borio, C., 2011. Rediscovering the macroeconomic roots of financial stability policy: journey, challenges and a way forward. *Annual Review of Financial Economics* 3, 87-117.
- Brunnermeier, M. and L. Pedersen, 2009, "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22 (6), 2201-2238.
- DeMarzo, P., M. Fishman, Z. He, and N. Wang, 2011, "Dynamic Agency and q Theory of Investment," *Journal of Finance*, forthcoming.
- Duffie, D., A. Li, and T. Luebkke, 2010, "Policy Perspectives on OTC Derivatives Market Infrastructure," MFI Working Paper 2010-002.
- Ellul, A. and V. Yerramilli, 2013, "Stronger Risk Controls, Lower Risk: Evidence from US Bank Holding Companies," *Journal of Finance*, 68 (5), 1757-1803.

- European Central Bank, 2009, "Credit Default Swaps and Counterparty Risk," August.
- Fender, I. and M. Scheicher, 2008, "The ABX: how do the markets price subprime mortgage risk?," *BIS Quarterly Review*, September, 67-81.
- Froot, K., Scharfstein, D. and J. Stein, 1993, "Risk Management: Coordinating Corporate Investment and Financing Policies," *Journal of Finance*, 48 (5), 1629-1658.
- Gorton, G., and G. Pennacchi, 1990, "Financial Intermediaries and Liquidity Creation," *Journal of Finance*, 45 (1), 49-71.
- Gromb D, and D. Vayanos, 2002, "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs," *Journal of Financial Economics*, 66 (2-3), 361-407.
- Holmström, B. and J. Tirole, 1998, "Private and Public Supply of Liquidity," *Journal of Political Economy*, 106 (1), 1-40.
- Jensen, M., and Meckling, 1976, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3 (4), 305-360.
- Lorenzoni, G., 2008, "Inefficient Credit Booms," *Review of Economic Studies*, 75 (3), 809-833.
- Myers, S., 1977, "Determinants of Corporate Borrowing," *Journal of Financial Economics*, 5 (2), 147-175.
- Myers, S. and R. Rajan, 1998, "The Paradox of Liquidity", *Quarterly Journal of Economics*, 113 (3), 733-771.
- Parlour, C. and G. Plantin, 2008, "Loan Sales and Relationship Banking," *Journal of Finance*, 63 (3), 1291-1314.
- Rajan, R., 2006, "Has Financial Development Made the World Riskier?", *European Financial Management*, 12 (4), 499-533.
- Rampini, A. and S. Viswanathan, 2010, "Collateral, Risk Management, and the Distribution of Debt Capacity," *Journal of Finance*, 65 (6), 2293-2322.
- Stephens, E. and J.R. Thompson, 2011, "CDS as Insurance: Leaky Lifeboats in Stormy Seas," working paper, University of Alberta and University of Waterloo.
- Thompson, J.R., 2010, "Counterparty Risk in Financial Contracts: Should the Insured Worry about the Insurer?," *Quarterly Journal of Economics*, 125 (3), 1195-1252.
- Tirole, J., 2006, *The Theory of Corporate Finance* (Princeton University Press, Prince-



ton).

## 8 Appendix

**Proof of Proposition 1** Form the Lagrangian using the objective (6), the feasibility constraints (3) with multiplier  $\mu_{FC}$  and the participation constraint (7) with multiplier  $\mu$ . For the moment we ignore the limited liability constraints (4) in the first-best. We then show that first-best transfers do not violate limited liability given our assumption  $AR > \pi\Delta\theta$ . Since  $\tilde{R} = R$  under effort, we do not explicitly write the dependence of the transfers on  $\tilde{R}$ .

The first-order conditions of the Lagrangian with respect to  $\tau^B(\theta, s)$  and  $\tau^S(\theta, s)$  are, respectively,

$$\text{prob}[\theta, s]u'(\theta + \tau^B(\theta, s)) - \mu_{FC}(\theta, s) = 0 \quad \forall(\theta, s) \quad (43)$$

$$\mu\text{prob}[\theta, s] - \mu_{FC}(\theta, s) = 0 \quad \forall(\theta, s). \quad (44)$$

Since marginal utility is strictly positive, it follows from (43) that  $\mu_{FC}(\theta, s) > 0$  for all  $(\theta, s)$  and hence the feasibility constraints bind. Since  $\mu_{FC}(\theta, s) > 0$ , it follows from (44) that the participation constraint binds. After substituting (43) into (44), it follows that buyers' marginal utility is the same across all states. That is, there is full risk-sharing.

From equal marginal utility across all states, we obtain, first, that  $\theta + \tau^B(\theta, \bar{s}) = \theta + \tau^B(\theta, \underline{s})$  and hence  $\tau^B(\theta, \bar{s}) = \tau^B(\theta, \underline{s})$  for  $\theta = \bar{\theta}, \underline{\theta}$ . Second, we obtain that  $\bar{\theta} + \tau^B(\bar{\theta}, s) = \underline{\theta} + \tau^B(\underline{\theta}, s)$  and hence  $\tau^B(\underline{\theta}, s) - \tau^B(\bar{\theta}, s) = \Delta\theta$  for  $s = \bar{s}, \underline{s}$ .

Using  $\tau^S(\theta, s) = -\tau^B(\theta, s)$  (from the binding feasibility constraints) and  $\tau^B(\theta, \bar{s}) = \tau^B(\theta, \underline{s})$ , we can write the binding participation constraint as

$$-(\text{prob}[\bar{\theta}, \bar{s}] + \text{prob}[\bar{\theta}, \underline{s}])\tau^B(\bar{\theta}, \bar{s}) - (\text{prob}[\underline{\theta}, \bar{s}] + \text{prob}[\underline{\theta}, \underline{s}])\tau^B(\underline{\theta}, \bar{s}) = 0 \quad (45)$$

Using  $\tau^B(\underline{\theta}, \bar{s}) - \tau^B(\bar{\theta}, \bar{s}) = \Delta\theta$  to substitute for  $\tau^B(\bar{\theta}, \bar{s})$  and since  $\text{prob}[\bar{\theta}, \bar{s}] + \text{prob}[\bar{\theta}, \underline{s}] = \text{prob}[\bar{\theta}] = \pi$  (and similarly for  $1 - \pi$ ), the binding participation constraint yields

$$\tau^B(\underline{\theta}, \bar{s}) = \pi\Delta\theta \quad (46)$$

from which the remaining transfers in the proposition follow immediately. QED

**Proof of Lemma 1** Plugging the first-best transfers from Proposition 1 into the incentive conditions (10) and using  $\alpha(\bar{s}) = 0$  yields  $A\mathcal{P} \geq (\pi - \bar{\pi})\Delta\theta$  and  $A\mathcal{P} \geq (\pi - \underline{\pi})\Delta\theta$ .

When the signal is informative,  $\lambda > \frac{1}{2}$ , we have  $\bar{\pi} > \pi > \underline{\pi}$ . The result in the lemma follows. QED

**Proof of Proposition 2** Form the Lagrangian using the objective (6), the feasibility constraints (3) with multiplier  $\mu_{FC}(\theta, s)$ , the limited liability constraints (4) with multipliers  $\mu_{LL}(\theta, s)$ , the feasibility constraints on margins (5) with  $\mu_0(s)$  for  $\alpha(s) \geq 0$  and  $\mu_1(s)$  for  $\alpha(s) \leq 1$ , the incentive compatibility constraints (10) with multipliers  $\mu_{IC}(s)$  and the participation constraint (12) with multiplier  $\mu$ .

The first-order conditions of the Lagrangian with respect to  $\tau^B(\theta, s)$  and  $\tau^S(\theta, s)$  are

$$\text{prob}[\theta, s]u'(\theta + \tau^B(\theta, s)) - \mu_{FC}(\theta, s) = 0 \quad \forall(\theta, s) \quad (47)$$

$$\mu \text{prob}[\theta, s] + \mu_{LL}(\theta, s) + \text{prob}[\theta|s]\mu_{IC}(s) - \mu_{FC}(\theta, s) = 0 \quad \forall(\theta, s). \quad (48)$$

Since marginal utilities are positive, it follows from (47) that  $\mu_{FC}(\theta, s) > 0$  and hence all feasibility constraints bind:

$$\tau^B(\theta, s) = -\tau^S(\theta, s), \forall(\theta, s). \quad (49)$$

Using (47) to substitute for  $\mu_{FC}(\theta, s)$  in (48) and rearranging, we obtain

$$u'(\theta + \tau^B(\theta, s)) = \mu + \frac{\mu_{LL}(\theta, s)}{\text{prob}[\theta, s]} + \frac{\mu_{IC}(s)}{\text{prob}[s]} \quad \forall(\theta, s) \quad (50)$$

where we used that  $\text{prob}[\theta|s]\text{prob}[s] = \text{prob}[\theta, s]$ .

We next show that the limited liability constraint in state  $(\bar{\theta}, s)$  is slack for each  $s$ . The proof proceeds in two steps. First, we show that the limited liability constraints cannot bind for both the state  $(\bar{\theta}, s)$  and the state  $(\underline{\theta}, s)$ . Suppose not. Since both limited liability constraints after the signal  $s$  bind, we have  $-\tau^S(\bar{\theta}, s) = \alpha(s)A + (1 - \alpha(s))AR$  and  $-\tau^S(\underline{\theta}, s) = \alpha(s)A + (1 - \alpha(s))AR$ . Hence,

$$E[-\tau^S(\bar{\theta}, \tilde{s})|\tilde{s} = s] = \alpha(s)A + (1 - \alpha(s))AR \quad \forall s$$

But since  $R > \mathcal{P}$ , this violates the incentive compatibility constraint (10) after the signal  $s$ . Hence, at least one limited liability constraint after the signal  $s$  must be slack.

Second, we show that the limited liability constraint in state  $(\bar{\theta}, s)$  is always slack for each  $s$ . Suppose not, so that  $-\tau^S(\bar{\theta}, s) = \alpha(s)A + (1 - \alpha(s))AR$ . We have just shown that at

least one limited liability constraint after the signal  $s$  must be slack. Hence, we must have that  $-\tau^S(\underline{\theta}, s) < \alpha(s)A + (1 - \alpha(s))AR$  and  $\mu_{LL}(\underline{\theta}, s) = 0$ . Using the binding feasibility constraints (49), we therefore have

$$\tau^B(\bar{\theta}, s) > \tau^B(\underline{\theta}, s) \quad \forall s,$$

which implies

$$u'(\bar{\theta} + \tau^B(\bar{\theta}, s)) < u'(\underline{\theta} + \tau^B(\underline{\theta}, s)) \quad \forall s$$

since  $\bar{\theta} > \underline{\theta}$ . However, using  $\mu_{LL}(\underline{\theta}, s) = 0$  in (50) implies

$$u'(\bar{\theta} + \tau^B(\bar{\theta}, s)) \geq u'(\underline{\theta} + \tau^B(\underline{\theta}, s)) \quad \forall s.$$

A contradiction. Hence, the limited liability constraint is slack in state  $(\bar{\theta}, s)$  and  $\mu_{LL}(\bar{\theta}, s) = 0$  for all  $s$ .

Finally, we show by contradiction that the participation constraint (12) binds. Suppose not. Plugging  $\mu = 0$  and  $\mu_{LL}(\bar{\theta}, s) = 0$  (just shown above) into (50) implies that  $\mu_{IC}(s) > 0$  for all  $s$ . Hence, both incentive constraints bind,  $-E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = s] = \alpha(s)A + (1 - \alpha(s))A\mathcal{P}$  for  $s = \bar{s}, \underline{s}$ . Therefore,

$$E[\tau^S(\tilde{\theta}, \tilde{s})] = E[E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s}]] = -E[\alpha(\tilde{s})A + (1 - \alpha(\tilde{s}))A\mathcal{P}] \quad (51)$$

From the participation constraint, we have

$$\begin{aligned} 0 &\leq E[\tau^S(\tilde{\theta}, \tilde{s})] - E[\alpha(\tilde{s})]A(R - C - 1) \\ &= -E[\alpha(\tilde{s})A + (1 - \alpha(\tilde{s}))A\mathcal{P}] - E[\alpha(\tilde{s})]A(R - C - 1) \quad [\text{using (51)}] \\ &= -E[(1 - \alpha(\tilde{s}))A\mathcal{P} + \alpha(\tilde{s})A(R - C)]. \end{aligned}$$

The last expression is strictly negative since  $R - C > \mathcal{P} > 0$  and  $0 \leq \alpha(\tilde{s}) \leq 1$ . A contradiction. Hence, the participation constraint binds and also  $\mu > 0$ . QED

### Proof of Proposition 3

The first-order conditions of the Lagrangian from the proof of Proposition 2 with respect to  $\alpha(s)$  are

$$\frac{\mu_0(s) - \mu_1(s)}{A} + \mu_{IC}(s)(1 - \mathcal{P}) = \mu \text{prob}[s](R - C - 1) + (R - 1)\mu_{LL}(\underline{\theta}, s) \quad \forall s, \quad (52)$$

where we have used  $\mu_{LL}(\bar{\theta}, s) = 0$  for all  $s$  (Proposition 2).

The right-hand side of (52) is strictly positive since  $R - C > 1$  and  $\mu > 0$  (see the end of the proof of Proposition 2). If the incentive constraint is slack for a signal  $s$ , then  $\mu_s = 0$ , implying that  $\mu_0(s) > 0$  must hold and  $\alpha(s) = 0$ . Similarly, if  $\mathcal{P} \geq 1$ , then  $\mu_0(s) > 0$  for each  $s$  must hold and  $\alpha(s) = 0$  for all  $s$ . QED

#### Proof of Proposition 4

It cannot be that both incentive constraints are slack since we assume that the first-best is not attainable,  $A\mathcal{P} < (\pi - \underline{\pi})\Delta\theta$ . It also cannot be that both incentive constraints bind (see the argument that the participation constraint binds in the proof of Proposition 2).

We now show by contradiction that the incentive constraint following a bad signal binds. Suppose not and hence  $\mu_{IC}(\underline{s}) = 0$ . After the good signal, the limited liability constraints are slack,  $\mu_{LL}(\bar{\theta}, \bar{s}) = 0$  by Proposition 2 and  $\mu_{LL}(\underline{\theta}, \bar{s}) = 0$  since we are considering a relaxed problem - see condition (14)). Equations (50) for  $s = \bar{s}$  then imply that

$$u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s})) = u'(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s})).$$

There is full risk-sharing conditional on the good signal. For transfers after a good signal we thus have

$$\tau^B(\underline{\theta}, \bar{s}) - \tau^B(\bar{\theta}, \bar{s}) = \Delta\theta > 0 \quad (53)$$

After the bad signal, limited liability constraint in state  $(\bar{\theta}, \underline{s})$  is slack,  $\mu_{LL}(\bar{\theta}, \underline{s}) = 0$  by Proposition 2. In state  $(\underline{\theta}, \underline{s})$ , we have two cases to consider, depending on whether the limited liability constraint is slack or whether it binds.

Consider first the case when the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is slack,  $\mu_{LL}(\underline{\theta}, \underline{s}) = 0$ . Equations (50) for  $s = \underline{s}$  then imply that there is also full risk-sharing conditional on the bad signal,

$$u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s})) = u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})).$$

and thus

$$\tau^B(\underline{\theta}, \underline{s}) - \tau^B(\bar{\theta}, \underline{s}) = \Delta\theta > 0 \quad (54)$$

Since  $\mu_{IC}(\underline{s}) = 0$  and  $\mu_{LL}(\underline{\theta}, \underline{s}) = \mu_{LL}(\underline{\theta}, \bar{s}) = 0$ , it follows from equations in (50) that

$$u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) \leq u'(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}))$$

and thus

$$\tau^B(\underline{\theta}, \underline{s}) \geq \tau^B(\underline{\theta}, \bar{s}). \quad (55)$$

From the binding participation constraint

$$\text{prob}[\bar{s}]E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}] + \text{prob}[\underline{s}]E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] = E[\alpha(\tilde{s})]A(R - C - 1) \geq 0$$

and  $E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}] < 0$  (binding incentive constraint after a good signal), we know that

$$E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] > 0 \quad (56)$$

Using full risk-sharing conditional on the signal (equations (53) and (54)) we can write

$$\begin{aligned} E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] &= \underline{\pi}\tau^S(\bar{\theta}, \underline{s}) + (1 - \underline{\pi})\tau^S(\underline{\theta}, \underline{s}) \\ &= \tau^S(\underline{\theta}, \underline{s}) + \underline{\pi}[\tau^S(\bar{\theta}, \underline{s}) - \tau^S(\underline{\theta}, \underline{s})] \\ &= \tau^S(\underline{\theta}, \underline{s}) + \underline{\pi}[\tau^S(\bar{\theta}, \bar{s}) - \tau^S(\underline{\theta}, \bar{s})] \end{aligned}$$

Using (55) and the binding feasibility conditions (49), we have  $\tau^S(\underline{\theta}, \underline{s}) \leq \tau^S(\underline{\theta}, \bar{s})$ . And since  $\underline{\pi} < \bar{\pi}$  (the signal is informative), we have

$$\begin{aligned} E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] &\leq \tau^S(\underline{\theta}, \bar{s}) + \underline{\pi}[\tau^S(\bar{\theta}, \bar{s}) - \tau^S(\underline{\theta}, \bar{s})] \\ &< \tau^S(\underline{\theta}, \bar{s}) + \bar{\pi}[\tau^S(\bar{\theta}, \bar{s}) - \tau^S(\underline{\theta}, \bar{s})] \end{aligned}$$

and thus  $E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] < E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]$ . But since  $E[\tau^S(\tilde{\theta}, \tilde{s})|\bar{s}] < 0$  (by the binding incentive constraint after a good signal), we have a contradiction with (56).

Now, consider the case when the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  binds. Since  $\mu_{LL}(\bar{\theta}, \underline{s}) = 0$  (by Proposition 2) and  $\mu_{IC}(\underline{s}) = 0$ , equations (50) for  $s = \underline{s}$  imply that

$$u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) \geq u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))$$

and thus

$$\tau^B(\underline{\theta}, \underline{s}) - \tau^B(\bar{\theta}, \underline{s}) \leq \Delta\theta. \quad (57)$$

Since  $\alpha(\underline{s}) = 0$  (incentive constraint after a bad signal is slack in contradiction), the binding limited liability constraint is  $AR = -\tau^S(\underline{\theta}, \underline{s})$ . Together with (57) in conjunction

with the binding feasibility constraints (49), we then have

$$\begin{aligned}
-E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] &= -[\underline{\pi}\tau^S(\bar{\theta}, \underline{s}) + (1 - \underline{\pi})\tau^S(\underline{\theta}, \underline{s})] \\
&= -\tau^S(\underline{\theta}, \underline{s}) - \underline{\pi}[\tau^S(\bar{\theta}, \underline{s}) - \tau^S(\underline{\theta}, \underline{s})] \\
&\geq AR - \underline{\pi}\Delta\theta
\end{aligned}$$

Since  $\underline{\pi} < \pi$  (informative signal) and  $AR > \pi\Delta\theta$  (limited liability constraints are slack in the first-best), we have  $-E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] > (\pi - \underline{\pi})\Delta\theta$ . But since the incentive constraint after a bad signal is slack,  $A\mathcal{P} > -E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]$ , this would mean that  $A\mathcal{P} > (\pi - \underline{\pi})\Delta\theta$  and the first-best can be reached, which is a contradiction.

Consequently, the incentive constraint after a bad signal binds and the incentive constraint after a good signal must be slack. QED

### Proof of Proposition 5

After a good signal, we have full risk-sharing (see the derivation of equation (53) in the proof of Proposition 4). Using (53) and (16), we obtain the transfers  $\tau^B(\bar{\theta}, \bar{s})$  and  $\tau^B(\underline{\theta}, \bar{s})$ .

After a bad signal, we have to distinguish two cases, depending on whether the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is slack or not. If it is slack, then we have full risk-sharing (see the derivation of equation (54) in the proof of Proposition 4). Using (54) and (15), we obtain the transfers  $\tau^B(\bar{\theta}, \underline{s})$  and  $\tau^B(\underline{\theta}, \underline{s})$ . If the limited liability constraint binds, we have  $\alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR = -\tau^S(\underline{\theta}, \underline{s})$ , which we plug into (15) to obtain  $\tau^B(\bar{\theta}, \underline{s})$ .

Finally, we check that, under (14), the limited liability constraint in  $(\underline{\theta}, \bar{s})$  is slack. Since  $\alpha(\bar{s}) = 0$ , the limited liability constraint (4) writes as  $\tau^B(\underline{\theta}, \bar{s}) < AR$ . Now, Proposition 5 implies that  $\tau^B(\underline{\theta}, \bar{s})$  decreases in  $\alpha(\underline{s})$ . So  $\tau^B(\underline{\theta}, \bar{s}) < AR$  for all  $\alpha(\underline{s})$  if and only if it is for  $\alpha(\underline{s}) = 0$ . After simplifications,  $\tau^B(\underline{\theta}, \bar{s}) < AR$  for  $\alpha(\underline{s}) = 0$  is equivalent to (14).

QED

### Proof of Proposition 6

We claim that  $\alpha^*(\underline{s}) < 1$ . Suppose not and  $\alpha^*(\underline{s}) = 1$ . First, note that  $\mu_{LL}(\underline{\theta}, \underline{s}) > 0$  must hold in this case. Suppose not, and  $\mu_{LL}(\underline{\theta}, \underline{s}) = 0$ . Then, equations (50) for  $s = \underline{s}$  imply that there is full risk-sharing conditional on the bad signal. Hence, the individual transfers after the bad signal are given by (18) so that  $\tau^B(\underline{\theta}, \underline{s}) = -\tau^S(\underline{\theta}, \underline{s}) = \underline{\pi}\Delta\theta + A > A$ . But the limited liability constraint requires  $-\tau^S(\underline{\theta}, \underline{s}) \leq A$ , a contradiction. Since  $\mu_{LL}(\underline{\theta}, \underline{s}) > 0$ , the

limited liability constraint binds and the individual transfers after a bad signal are as in (19). In particular,  $\tau^B(\bar{\theta}, \underline{s}) = A > 0$ . Equations (50) and binding incentive constraint after a bad signal imply that  $\tau^B(\bar{\theta}, \bar{s}) \geq \tau^B(\bar{\theta}, \underline{s}) = A > 0$ . However, by equation (17),  $\tau^B(\bar{\theta}, \bar{s}) < 0$ . A contradiction.

QED

### Proof of Proposition 7

Since the incentive constraint after a good signal is slack (see Proposition 4), it follows from Proposition 3 that  $\alpha^*(\bar{s}) = 0$ . It remains to characterize the optimal margin after a bad signal.

We now derive the optimal margin after a bad signal,  $\alpha^*(\underline{s})$ . Using equations (50) to substitute for  $\mu$ ,  $\mu_{IC}(\underline{s})$  and  $\mu_{LL}(\underline{\theta}, \underline{s})$  in equation (52), we get

$$\begin{aligned} \frac{u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))} &= 1 + \frac{R - C - 1}{1 - \mathcal{P}} + \frac{\mu_1(\underline{s}) - \mu_0(\underline{s})}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))\text{prob}[\underline{s}](1 - \mathcal{P})A} \\ &\quad + \frac{1 - \pi}{1 - \mathcal{P}} \frac{u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) - u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))} \end{aligned} \quad (58)$$

where we used  $\mu_{IC}(\bar{s}) = 0$  (Proposition 4).

Denote the RHS of (58) by  $\varphi$ . Note that  $\frac{\partial \tau^B(\bar{\theta}, \bar{s})}{\partial \alpha} = -\frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A(R - C - \mathcal{P}) < 0$ . For  $\mathcal{P} < 1$ ,  $\frac{\partial \tau^B(\bar{\theta}, \underline{s})}{\partial \alpha} > 0$ . (When the limited liability constraint is slack, we have  $\frac{\partial \tau^B(\bar{\theta}, \underline{s})}{\partial \alpha} = A(1 - \mathcal{P}) > 0$  and when the limited liability constraint binds, we have  $\frac{\partial \tau^B(\bar{\theta}, \underline{s})}{\partial \alpha} = A\left[1 + \frac{(1 - \pi)R - \mathcal{P}}{\pi}\right] > 0$  since  $R - \mathcal{P} > R - 1 > \pi(R - 1)$ ). Hence,  $\varphi$  is decreasing in  $\alpha$ . If  $\varphi(0) < 1 + \frac{R - C - 1}{1 - \mathcal{P}}$ , then

$$\varphi(0) < 1 + \frac{R - C - 1}{1 - \mathcal{P}} + \frac{1 - \pi}{1 - \mathcal{P}} \frac{u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) - u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))}$$

for any  $\alpha \in [0, 1]$  (since the last term is non-negative). By equation (58) we have  $\mu_0 > 0$  and hence  $\alpha^*(\underline{s}) = 0$ .

Otherwise, there are two cases depending on whether or not the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is slack. If it is slack, then marginal utilities after the bad signal are equalized (equation (54)), and the last term in equation (58) vanishes. The optimal margin  $\alpha^*(\underline{s}) \in (0, 1)$  is given by  $\varphi(\alpha^*(\underline{s})) = 1 + \frac{R - C - 1}{1 - \mathcal{P}}$  in this case. If the limited liability constraint binds, then the optimal margin  $\alpha^*(\underline{s}) \in (0, 1)$  solves

$$\frac{u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))} - \frac{1 - \pi}{1 - \mathcal{P}} \frac{u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) - u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))} = 1 + \frac{R - C - 1}{1 - \mathcal{P}}$$



We now check under what conditions the limited liability constraints are slack. By Proposition 2, we only need to check limited liability constraints in states  $(\underline{\theta}, \bar{s})$  and  $(\underline{\theta}, \underline{s})$ . First, consider the case when  $\mathcal{P} \geq 1$  and margins are not used. The limited liability constraints are slack if and only if:

$$AR > -\tau^S(\theta, s, R) = \tau^B(\theta, s, R), \quad \forall(\theta, s, R)$$

Since  $\tau^B(\underline{\theta}, \bar{s}) \geq \tau^B(\underline{\theta}, \underline{s})$ , we only need to check when the limited liability constraint is slack in state  $(\underline{\theta}, \bar{s})$ . It is slack if and only if:

$$\bar{\pi}\Delta\theta - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A\mathcal{P} < AR \quad (59)$$

or, equivalently,

$$AR - \pi\Delta\theta > \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]} [(\pi - \bar{\pi})\Delta\theta - A\mathcal{P}] > 0.$$

Now consider the case when  $\mathcal{P} < 1$ . The limited liability constraints in this case are slack if and only if:

$$\alpha(s)A + (1 - \alpha(s))AR > -\tau^S(\theta, s, R) \quad \forall(\theta, s, R)$$

with  $\alpha^*(\bar{s}) = 0$  and  $\alpha^*(\underline{s}) \geq 0$ . The limited liability constraint in state  $(\underline{\theta}, \bar{s})$  is slack if and only if:

$$\bar{\pi}\Delta\theta - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A[\alpha^*(\underline{s})(R - C) + (1 - \alpha^*(\underline{s}))\mathcal{P}] < AR$$

Since  $R - C > \mathcal{P} > 0$ , we have:

$$\bar{\pi}\Delta\theta - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A[\alpha^*(\underline{s})(R - C) + (1 - \alpha^*(\underline{s}))\mathcal{P}] < \bar{\pi}\Delta\theta - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A\mathcal{P}$$

Hence, condition (59) is sufficient for the limited liability constraint to be slack in state  $(\underline{\theta}, \bar{s})$ .

The limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is slack if and only if:

$$\alpha^*(\underline{s}) < 1 - \frac{\pi\Delta\theta}{A(R - \mathcal{P})}$$

Since the optimal interior margin when the limited liability constraint is slack is given by

$$\alpha^*(\underline{s}) = \varphi^{-1}\left(1 + \frac{R - C - 1}{1 - \mathcal{P}}\right),$$

the constraint in state  $(\underline{\theta}, \underline{s})$  is slack if and only if

$$\varphi^{-1} \left( 1 + \frac{R - C - 1}{1 - \mathcal{P}} \right) < 1 - \frac{\pi \Delta \theta}{A(R - \mathcal{P})}.$$

Note that if the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is slack, it must be that  $\tau^B(\bar{\theta}, \underline{s}) < 0$  (equation (18)) implying that

$$\alpha^*(\underline{s}) < \frac{(1 - \pi)\Delta\theta - A\mathcal{P}}{A(1 - \mathcal{P})}$$

must hold if the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is slack.

In case the limited liability constraint binds, it also must be that  $\tau^B(\bar{\theta}, \underline{s}) < 0$ . This is because equations (19) imply that

$$\begin{aligned} \tau^B(\underline{\theta}, \underline{s}) &= \alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR > \\ E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] &= \alpha(\underline{s})A + (1 - \alpha(\underline{s}))A\mathcal{P} \quad [\text{since } R > \mathcal{P} \text{ and } \alpha^*(\underline{s}) < 1] \\ &> 0 > \tau^B(\bar{\theta}, \underline{s}) \quad [\text{since } E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] = \pi\tau^B(\bar{\theta}, \underline{s}) + (1 - \pi)\tau^B(\underline{\theta}, \underline{s})] \end{aligned}$$

For  $\tau^B(\bar{\theta}, \underline{s})$  to be negative if the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  binds, it must be that

$$\alpha^*(\underline{s}) \left[ 1 + \frac{(1 - \pi)R - \mathcal{P}}{\pi} \right] < \frac{(1 - \pi)R - \mathcal{P}}{\pi}$$

or, equivalently,

$$\alpha^*(\underline{s}) < \frac{(1 - \pi)R - \mathcal{P}}{\pi + (1 - \pi)R - \mathcal{P}} < 1$$

It follows that a sufficient condition for the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  to be slack is

$$1 - \frac{\pi \Delta \theta}{A(R - \mathcal{P})} > \frac{(1 - \pi)R - \mathcal{P}}{\pi + (1 - \pi)R - \mathcal{P}}.$$

QED

### Proof of Proposition 8

Form the Lagrangian using the objective (27), the feasibility constraints (28) and (29) with multipliers  $\mu_{FC}(\theta, s)$ , the limited liability constraints (30) and (31) with multipliers  $\mu_{LL}(\theta, s, R)$ , the feasibility constraints on margins (5) with  $\mu_0(s)$  for  $\alpha(s) \geq 0$  and  $\mu_1(s)$  for  $\alpha(s) \leq 1$ , the incentive compatibility constraints (32) and (33) with multipliers  $\mu_{IC}(s)$  and the participation constraint (34) with multiplier  $\mu$ .

The first-order conditions of the Lagrangian with respect to  $\tau^B(\theta, s)$  are

$$\text{prob}[\theta, s]u'(\theta + \tau^B(\theta, s)) - \mu_{FC}(\theta, s) = 0 \quad \forall(\theta, s) \quad (60)$$

The first-order conditions of the Lagrangian with respect to  $\tau^S(\theta, \bar{s}, R)$ ,  $\tau^S(\theta, \underline{s}, R)$  and  $\tau^S(\theta, \underline{s}, 0)$  are

$$\mu \text{prob}[\theta, \bar{s}] + \mu_{LL}(\theta, \bar{s}, R) + \text{prob}[\theta|\bar{s}]\mu_{IC}(\bar{s}) - \mu_{FC}(\theta, \bar{s}) = 0 \quad \forall(\theta, \bar{s}, R) \quad (61)$$

$$\mu \text{prob}[\theta, \underline{s}] + \frac{\mu_{LL}(\theta, \underline{s}, R)}{p} - \text{prob}[\theta|\underline{s}]\frac{\mu_{IC}(\underline{s})}{p} - \mu_{FC}(\theta, \underline{s}) = 0 \quad \forall(\theta, \underline{s}, R) \quad (62)$$

$$\mu \text{prob}[\theta, \underline{s}] + \frac{\mu_{LL}(\theta, \underline{s}, 0)}{1-p} + \text{prob}[\theta|\underline{s}]\frac{\mu_{IC}(\underline{s})}{1-p} - \mu_{FC}(\theta, \underline{s}) = 0 \quad \forall(\theta, \underline{s}, 0) \quad (63)$$

Since marginal utilities are positive, it follows from (60) that  $\mu_{FC}(\theta, s) > 0$  and hence the feasibility constraints (28) and (29) bind.

Using (60) to substitute for  $\mu_{FC}(\theta, s)$  in (61)-(63) and rearranging, we obtain

$$u'(\theta + \tau^B(\theta, \bar{s})) = \mu + \frac{\mu_{LL}(\theta, \bar{s}, R)}{\text{prob}[\theta, \bar{s}]} + \frac{\mu_{IC}(\bar{s})}{\text{prob}[\bar{s}]} \quad \forall(\theta, \bar{s}, R) \quad (64)$$

$$u'(\theta + \tau^B(\theta, \underline{s})) = \mu + \frac{\mu_{LL}(\theta, \underline{s}, R)}{p \text{prob}[\theta, \underline{s}]} - \frac{\mu_{IC}(\underline{s})}{p \text{prob}[\underline{s}]} \quad \forall(\theta, \underline{s}, R) \quad (65)$$

$$u'(\theta + \tau^B(\theta, \underline{s})) = \mu + \frac{\mu_{LL}(\theta, \underline{s}, 0)}{(1-p) \text{prob}[\theta, \underline{s}]} + \frac{\mu_{IC}(\underline{s})}{(1-p) \text{prob}[\underline{s}]} \quad \forall(\theta, \underline{s}, 0) \quad (66)$$

where we used that  $\text{prob}[\theta|s]\text{prob}[s] = \text{prob}[\theta, s]$ .

Combining (65) and (66) yields

$$(1-p)\mu_{LL}(\theta, \underline{s}, R) - p\mu_{LL}(\theta, \underline{s}, 0) = \text{prob}[\theta|\underline{s}]\mu_{IC}(\underline{s}) \quad \forall(\theta, \underline{s}) \quad (67)$$

We next show that the limited liability constraint in state  $(\bar{\theta}, \bar{s}, R)$  is slack. The proof proceeds in two steps. First, we show that the limited liability constraints cannot bind for both the state  $(\bar{\theta}, \bar{s}, R)$  and the state  $(\underline{\theta}, \bar{s}, R)$ . Suppose not. Since both limited liability constraints after the signal  $\bar{s}$  bind, we have  $-\tau^S(\bar{\theta}, \bar{s}, R) = \alpha(\bar{s})A + (1 - \alpha(\bar{s}))AR$  and  $-\tau^S(\underline{\theta}, \bar{s}, R) = \alpha(\bar{s})A + (1 - \alpha(\bar{s}))AR$ . Hence,

$$E[-\tau^S(\theta, \bar{s}, R)] = \alpha(\bar{s})A + (1 - \alpha(\bar{s}))AR$$

But since  $R > \mathcal{P}$ , this violates the incentive compatibility constraint (32) after the good signal. Hence, at least one limited liability constraint after the signal  $\bar{s}$  must be slack.

Second, we show that the limited liability constraint in state  $(\bar{\theta}, \bar{s}, R)$  is always slack. Suppose not, so that  $-\tau^S(\bar{\theta}, \bar{s}, R) = \alpha(\bar{s})A + (1 - \alpha(\bar{s}))AR$ . We have just shown that at least one limited liability constraint after the signal  $\bar{s}$  must be slack. Hence, we must have that  $-\tau^S(\underline{\theta}, \bar{s}, R) < \alpha(\bar{s})A + (1 - \alpha(\bar{s}))AR$  and  $\mu_{LL}(\underline{\theta}, \bar{s}, R) = 0$ . Using the binding feasibility constraints (28), we have

$$\tau^B(\bar{\theta}, \bar{s}, R) > \tau^B(\underline{\theta}, \bar{s}, R),$$

which implies

$$u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}, R)) < u'(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}, R))$$

since  $\bar{\theta} > \underline{\theta}$ . However, using  $\mu_{LL}(\underline{\theta}, \bar{s}, R) = 0$  in (64) implies

$$u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}, R)) \geq u'(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}, R)).$$

A contradiction. Hence, the limited liability constraint is slack in state  $(\bar{\theta}, \bar{s}, R)$  and  $\mu_{LL}(\bar{\theta}, \bar{s}, R) = 0$ .

Third, we show by contradiction that  $\mu > 0$  and the participation constraint (34) binds. Suppose not, i.e.  $\mu = 0$ . Using  $\mu = 0$  in (65), it follows that  $\mu_{LL}(\theta, \underline{s}, R) > 0$  must hold for  $\theta = \bar{\theta}, \underline{\theta}$ . Using  $\mu = 0$  and  $\mu_{LL}(\bar{\theta}, \bar{s}, R) = 0$  (just shown above) in (64), it follows that  $\mu_{IC}(\bar{s}) > 0$  and the incentive constraint in state  $\bar{s}$  binds. Now, there are two possibilities in state  $\underline{s}$ : either the incentive constraint binds or it is slack.

Consider first the case when the incentive constraint in state  $\underline{s}$  binds. Using the binding limited liability constraints in states  $(\bar{\theta}, \underline{s}, R)$  and  $(\underline{\theta}, \underline{s}, R)$  in the incentive constraint in state  $\underline{s}$ , we get

$$(1 - \alpha(\underline{s}))A\mathcal{P} = \alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR + \underline{\pi}\tau^S(\bar{\theta}, \underline{s}, 0) + (1 - \underline{\pi})\tau^S(\underline{\theta}, \underline{s}, 0)$$

or, equivalently,

$$\alpha(\underline{s})A + (1 - \alpha(\underline{s}))A(R - \mathcal{P}) = -\underline{\pi}\tau^S(\bar{\theta}, \underline{s}, 0) - (1 - \underline{\pi})\tau^S(\underline{\theta}, \underline{s}, 0) \quad (68)$$

If the limited liability constraints (31) are slack, we have  $-\tau^S(\bar{\theta}, \underline{s}, 0) < \alpha(\underline{s})A$  and  $-\tau^S(\underline{\theta}, \underline{s}, 0) < \alpha(\underline{s})A$  so that the right-hand side of (68) is strictly smaller than  $\alpha(\underline{s})A$ . Since  $(1 - \alpha(\underline{s}))A(R - \mathcal{P}) \geq 0$ , the left-hand side of (68) is greater or equal to  $\alpha(\underline{s})A$ . A contradiction. If the limited liability constraints (31) are binding, then all limited liability constraints in state  $\underline{s}$  bind.

Using the binding limited liability constraints in state  $\underline{s}$  and the binding incentive constraint in state  $\bar{s}$  in the (weakly slack) participation constraint (34), we get

$$\begin{aligned} & \text{prob}[\bar{s}]\alpha(\bar{s})A(R - C - 1) + \text{prob}[\underline{s}]\alpha(\underline{s})A(pR - 1) + \text{prob}[\underline{s}](1 - p)A\mathcal{P} \\ & \leq -\text{prob}[\bar{s}](\alpha(\bar{s})A + (1 - \alpha(\bar{s}))A\mathcal{P}) - \text{prob}[\underline{s}](p(\alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR) + (1 - p)\alpha(\underline{s})A) \end{aligned}$$

Simplifying yields

$$\text{prob}[\bar{s}][\alpha(\bar{s})A(R - C) + (1 - \alpha(\bar{s}))A\mathcal{P}] + \text{prob}[\underline{s}]A[(1 - p)\mathcal{P} + pR] \leq 0 \quad (69)$$

Since both terms on the right-hand side of (69) are strictly positive, we have a contradiction.

Now consider the case when the incentive constraint in state  $\underline{s}$  is slack so that  $\mu_{IC}(\underline{s}) = 0$ . Since  $\mu_{LL}(\bar{\theta}, \underline{s}, R) > 0$  and  $\mu_{LL}(\underline{\theta}, \underline{s}, R) > 0$ , using  $\mu_{IC}(\underline{s}) = 0$  in (67) implies that  $\mu_{LL}(\bar{\theta}, \underline{s}, 0) > 0$  and  $\mu_{LL}(\underline{\theta}, \underline{s}, 0) > 0$  must hold. Hence, all limited liability constraints in state  $\underline{s}$  bind. But we have just shown in the previous step that this is incompatible with the weakly slack participation constraint. A contradiction.

We conclude that  $\mu > 0$  and the participation constraint must bind.

Fourth, we show that  $\mu_{LL}(\bar{\theta}, \underline{s}, R) = 0$  and  $-\tau^S(\bar{\theta}, \underline{s}, R) \leq \alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR$ . The proof proceeds in two steps. First, we show that it cannot be that both  $\mu_{LL}(\bar{\theta}, \underline{s}, R) > 0$  and  $\mu_{LL}(\underline{\theta}, \underline{s}, R) > 0$ . Suppose not. When both  $\mu_{LL}(\bar{\theta}, \underline{s}, R) > 0$  and  $\mu_{LL}(\underline{\theta}, \underline{s}, R) > 0$ , then

$$-\tau^S(\bar{\theta}, \underline{s}, R) = -\tau^S(\underline{\theta}, \underline{s}, R) = \alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR \quad (70)$$

Using (70) in the incentive constraint after a bad signal (33) yields

$$-E[\tau^S(\theta, \underline{s}, 0)] + (1 - \alpha(\underline{s}))A\mathcal{P} < \alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR$$

since  $-E[\tau^S(\theta, \underline{s}, 0)] \leq \alpha(\underline{s})A$  and  $\mathcal{P} < R$ . Hence, the incentive constraint after a bad signal is slack and  $\mu_{IC}(\underline{s}) = 0$ . Since  $\mu_{LL}(\bar{\theta}, \underline{s}, R) > 0$  and  $\mu_{LL}(\underline{\theta}, \underline{s}, R) > 0$ , using  $\mu_{IC}(\underline{s}) = 0$  in (67) implies that  $\mu_{LL}(\bar{\theta}, \underline{s}, 0) > 0$  and  $\mu_{LL}(\underline{\theta}, \underline{s}, 0) > 0$  must hold. Hence, all limited liability constraints in state  $\underline{s}$  bind. Using the binding limited liability constraints in state  $\underline{s}$  in the binding participation constraint (34), we get

$$\begin{aligned} & \text{prob}[\bar{s}]\alpha(\bar{s})A(R - C - 1) + \text{prob}[\underline{s}]\alpha(\underline{s})A(pR - 1) + \text{prob}[\underline{s}](1 - p)A\mathcal{P} \\ & = \text{prob}[\bar{s}]E[\tau^S(\theta, \bar{s}, R)] - \text{prob}[\underline{s}](p(\alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR) + (1 - p)\alpha(\underline{s})A) \end{aligned}$$

Simplifying yields

$$\text{prob}[\bar{s}]\alpha(\bar{s})A(R - C - 1) + \text{prob}[\underline{s}]ApR + \text{prob}[\underline{s}](1 - p)A\mathcal{P} = \text{prob}[\bar{s}]E[\tau^S(\theta, \bar{s}, R)] \quad (71)$$

For equation (71) to hold, it must be that  $E[\tau^S(\theta, \bar{s}, R)] > 0$ . By the binding feasibility constraint (28), this is equivalent to  $E[\tau^B(\theta, \bar{s}, R)] < 0$ . There can be two cases: either the incentive constraint after a good signal binds or it is slack. First, consider the case when the incentive constraint after a good signal binds. Then,  $E[\tau^S(\theta, \bar{s}, R)] = -(\alpha(\bar{s})A + (1 - \alpha(\bar{s}))A\mathcal{P}) < 0$ . A contradiction with (71). Second, consider the case when the incentive constraint after a good signal is slack. Then,  $\mu_{IC}(\bar{s}) = 0$ . Using  $\mu_{LL}(\theta, \bar{s}, R) = 0$  and  $\mu_{IC}(\bar{s}) = 0$  in (64) and  $\mu_{LL}(\bar{\theta}, \underline{s}, R) > 0$  and  $\mu_{IC}(\underline{s}) = 0$  in (65), we have

$$u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s})) < u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))$$

implying that  $\tau^B(\bar{\theta}, \bar{s}) > \tau^B(\bar{\theta}, \underline{s})$ . So, we have:

$$\begin{aligned} \tau^B(\bar{\theta}, \bar{s}) &> \tau^B(\bar{\theta}, \underline{s}) = -p\tau^S(\bar{\theta}, \underline{s}, R) - (1 - p)\tau^S(\bar{\theta}, \underline{s}, 0) && \text{[using binding feasibility constraint]} \\ &= p[\alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR] + (1 - p)\alpha(\underline{s})A && \text{[using binding LL constraints in state } \underline{s}] \\ &= \alpha(\underline{s})A + p(1 - \alpha(\underline{s}))AR > 0 && (72) \end{aligned}$$

Now, there are two cases to consider: either the limited liability constraint in state  $(\underline{\theta}, \bar{s}, R)$  binds or it is slack. If it binds, then  $\tau^B(\underline{\theta}, \bar{s}) = \alpha(\bar{s})A + (1 - \alpha(\bar{s}))AR > 0$ . Together with (72), this implies that  $E[\tau^B(\theta, \bar{s}, R)] > 0$ , a contradiction with (71). If the limited liability constraint in state  $(\underline{\theta}, \bar{s}, R)$  is slack, then  $\mu(\underline{\theta}, \bar{s}, R) = 0$ . Then, there is full risk-sharing after a good signal,  $\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}, R) = \underline{\theta} + \tau^B(\underline{\theta}, \bar{s}, R)$ , and  $\tau^B(\underline{\theta}, \bar{s}, R) = \tau^B(\bar{\theta}, \bar{s}, R) + \Delta\theta > 0$ . Together with (72), this implies that  $E[\tau^B(\theta, \bar{s}, R)] = -E[\tau^S(\theta, \bar{s}, R)] > 0$ , a contradiction with (71).

Hence, we showed that at least one of the  $\mu_{LL}(\theta, \underline{s}, R)$ 's must be zero. We now show that it is  $\mu_{LL}$  in state  $(\bar{\theta}, \underline{s}, R)$ . Suppose not, i.e.,  $\mu_{LL}(\bar{\theta}, \underline{s}, R) > 0$  and  $\mu_{LL}(\underline{\theta}, \underline{s}, R) = 0$ . Using  $\mu_{LL}(\underline{\theta}, \underline{s}, R) = 0$  in (67), it follows that  $\mu_{LL}(\underline{\theta}, \underline{s}, 0) = 0$  and  $\mu_{IC}(\underline{s}) = 0$ . Using  $\mu_{LL}(\bar{\theta}, \underline{s}, R) > 0$  and  $\mu_{IC}(\underline{s}) = 0$  in (67), it follows that  $\mu_{LL}(\bar{\theta}, \underline{s}, 0) > 0$ . Hence,

$$\tau^B(\bar{\theta}, \underline{s}) = p(\alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR) + (1 - p)\alpha(\underline{s})A \quad (73)$$

Using  $\mu_{LL}(\bar{\theta}, \underline{s}, R) > 0$  and  $\mu_{LL}(\underline{\theta}, \underline{s}, R) = 0$  in (65), we have

$$u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s})) > u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s}))$$

implying that  $\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}) < \underline{\theta} + \tau^B(\underline{\theta}, \underline{s})$ . Since  $\bar{\theta} > \underline{\theta}$ , this means that

$$\tau^B(\bar{\theta}, \underline{s}) < \tau^B(\underline{\theta}, \underline{s}) \tag{74}$$

must hold. However, we also have that

$$\begin{aligned} \tau^B(\underline{\theta}, \underline{s}) &= -p\tau^S(\underline{\theta}, \underline{s}, R) - (1-p)\tau^S(\underline{\theta}, \underline{s}, 0) \quad [\text{using binding feasibility constraint}] \\ &\leq p(\alpha(\underline{s})A + (1-\alpha(\underline{s}))AR) + (1-p)\alpha(\underline{s})A \quad [\text{using limited liability constraints}] \\ &= \tau^B(\bar{\theta}, \underline{s}) \quad [\text{using (73)}] \end{aligned}$$

which contradicts (74). Hence, we must have that  $\mu_{LL}(\bar{\theta}, \underline{s}, R) = 0$ .

Fifth, we claim that  $\mu_{LL}(\bar{\theta}, \underline{s}, 0) = 0$  and  $\mu_{IC}(\underline{s}) = 0$ . This claim follows immediately from substituting  $\mu_{LL}(\bar{\theta}, \underline{s}, R) = 0$  in (67).

QED

### Proof of Proposition 9

The first-order conditions of the Lagrangian from the proof of Proposition 8 with respect to  $\alpha(\bar{s})$  and  $\alpha(\underline{s})$  are

$$\frac{\mu_0(\bar{s}) - \mu_1(\bar{s})}{A} + \mu_{IC}(\bar{s})(1 - \mathcal{P}) = \mu\text{prob}[\bar{s}](R - C - 1) + (R - 1)\mu_{LL}(\underline{\theta}, \bar{s}, R) \tag{75}$$

$$\mu_{LL}(\underline{\theta}, \underline{s}, 0) + \frac{\mu_0(\underline{s}) - \mu_1(\underline{s})}{A} = \mu\text{prob}[\underline{s}](pR - 1) + (R - 1)\mu_{LL}(\underline{\theta}, \underline{s}, R) \tag{76}$$

where we have used  $\mu_{LL}(\bar{\theta}, \bar{s}, R) = 0$ ,  $\mu_{LL}(\bar{\theta}, \underline{s}, R) = 0$ ,  $\mu_{LL}(\bar{\theta}, \underline{s}, 0) = 0$  and  $\mu_{IC}(\underline{s}) = 0$  (all shown in the previous Proposition).

Consider first state  $\bar{s}$ . The right-hand side of (75) is strictly positive since  $R - C > 1$  and  $\mu > 0$  (see Proposition 8). If the incentive constraint is slack after a good signal, then  $\mu_{IC}(\bar{s}) = 0$ , implying that  $\mu_0(\bar{s}) > 0$  must hold and  $\alpha^*(\bar{s}) = 0$ . Similarly, if  $\mathcal{P} \geq 1$ , then  $\mu_0(\bar{s}) > 0$  must hold and  $\alpha^*(\bar{s}) = 0$ .

Consider now state  $\underline{s}$ . Using  $\mu_{IC}(\underline{s}) = 0$  (as shown in the previous Proposition) in (67) yields

$$\mu_{LL}(\underline{\theta}, \underline{s}, 0) = \frac{1-p}{p}\mu_{LL}(\underline{\theta}, \underline{s}, R)$$

Substituting for  $\mu_{LL}(\underline{\theta}, \underline{s}, 0)$  in (76) yields

$$\frac{\mu_0(\underline{s}) - \mu_1(\underline{s})}{A} = (pR - 1) \left[ \mu_{\text{prob}}[\underline{s}] + \frac{\mu_{LL}(\underline{\theta}, \underline{s}, R)}{p} \right] \quad (77)$$

If  $pR \geq 1$ , then the right-hand side of (77) is non-negative, implying that  $\mu_0(\underline{s}) \geq 0$  and  $\alpha^*(\underline{s}) = 0$ . If  $pR < 1$ , then the right-hand side of (77) is negative, implying that  $\mu_1(\underline{s}) > 0$  and  $\alpha^*(\underline{s}) = 1$ . We now claim that the contract with risk-taking and  $\alpha^*(\underline{s}) = 1$  is dominated by the contract with effort after a bad signal. Note that  $\alpha(\underline{s}) = 1$  is also feasible under the contract with effort. However, it is never chosen (Proposition 6), implying that the optimal contract with effort is strictly preferred to the contract with risk-taking and  $\alpha(\underline{s}) = 1$ .

QED

### Proof of Proposition 10

The optimal transfers follow from asserting full risk-sharing across all states and using the binding participation constraint. Condition (36) follows from checking that all limited liability constraints are satisfied for these transfers. It remains to check that, in the proposed contract, the incentive constraint after a good signal is slack and margins are not used. Using  $\alpha(\bar{s}) = 0$  and the transfers in state  $\bar{s}$  in the incentive constraint (32) we have:

$$A\mathcal{P} > 0 > -(\bar{\pi} - \pi)\Delta\theta - \text{prob}[\underline{s}](1-p)A\mathcal{P} = E[\tau^B(\theta, \bar{s}, R)] = -E[\tau^S(\theta, \bar{s}, R)]$$

so that the incentive constraint after  $\bar{s}$  is indeed slack at  $\alpha(\bar{s}) = 0$ . Since  $pR \geq 1$ , it is not optimal to use margins after a bad signal either (Proposition 9).

QED

### Proof of Proposition 11

We first show that for  $p < \max\{\frac{R-C-1}{R-1}, \frac{1}{R}\}$  the contract with effort is optimal. First, consider  $p \leq \frac{R-C-1}{R-1}$ . In this case, we have that  $\mathcal{P} \geq 1$ . Combining with condition (37) yields

$$A\mathcal{P} \geq A \geq \pi\Delta\theta > (\pi - \underline{\pi})\Delta\theta$$

By Lemma 1, the first-best (which entails effort) is reached. Second, consider  $p < \frac{1}{R}$ . By Proposition 9, the contract with effort strictly dominates the contract with risk-taking in this case.

We now consider the case when  $p \geq \max\{\frac{R-C-1}{R-1}, \frac{1}{R}\}$ . Note that  $p$  must always be lower than  $\frac{R-C}{R}$  since we require that  $\mathcal{P} > 0$ .



We now show that the expected utility of the contract with effort is decreasing in  $p$ . Consider first the case when the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is slack. Then, there is full risk-sharing conditional on the signal and, using Proposition 5, the expected utility of the protection buyer under effort is given by

$$\text{prob}[\bar{s}]u\left(E[\tilde{\theta}|\bar{s}] - \frac{\text{prob}[\underline{s}]A[\alpha^*(\underline{s})(R-C) + (1-\alpha^*(\underline{s}))\mathcal{P}]}{\text{prob}[\bar{s}]}\right) + \text{prob}[\underline{s}]u\left(E[\tilde{\theta}|\underline{s}] + A[\alpha^*(\underline{s}) + (1-\alpha^*(\underline{s}))\mathcal{P}]\right)$$

The derivative of the expected utility with respect to  $p$  is given by

$$\begin{aligned} & -\text{prob}[\bar{s}]u'\left(E[\tilde{\theta}|\bar{s}] - \frac{\text{prob}[\underline{s}]A[\alpha^*(\underline{s})(R-C) + (1-\alpha^*(\underline{s}))\mathcal{P}]}{\text{prob}[\bar{s}]}\right) \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A(1-\alpha^*(\underline{s}))\frac{\partial\mathcal{P}}{\partial p} + \\ & \text{prob}[\underline{s}]u'\left(E[\tilde{\theta}|\underline{s}] + A[\alpha^*(\underline{s}) + (1-\alpha^*(\underline{s}))\mathcal{P}]\right)A(1-\alpha^*(\underline{s}))\frac{\partial\mathcal{P}}{\partial p} = \text{prob}[\underline{s}]A(1-\alpha^*(\underline{s}))\frac{\partial\mathcal{P}}{\partial p} \times \\ & \left[ u'\left(E[\tilde{\theta}|\underline{s}] + A[\alpha^*(\underline{s}) + (1-\alpha^*(\underline{s}))\mathcal{P}]\right) - u'\left(E[\tilde{\theta}|\bar{s}] - \frac{\text{prob}[\underline{s}]A[\alpha^*(\underline{s})(R-C) + (1-\alpha^*(\underline{s}))\mathcal{P}]}{\text{prob}[\bar{s}]}\right) \right] \end{aligned}$$

where we have used the envelope theorem to claim  $\frac{\partial\alpha^*(\underline{s})}{\partial p} = 0$ . We know that  $1 - \alpha^*(\underline{s}) > 0$  since  $\alpha^*(\underline{s}) < 1$  (Proposition 6). Due to the binding incentive constraint after a bad signal (Proposition 4), the protection buyer's consumption is larger after a good signal than after a bad signal implying that the term in the square brackets above is positive. Since  $\mathcal{P} = R - \frac{C}{1-p}$ , we have  $\frac{\partial\mathcal{P}}{\partial p} < 0$  implying that the expected utility under effort decreases in  $p$  when the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is slack.

Now consider the other possibility, i.e., that the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is binding. Then, there is still full risk-sharing conditional on a good signal but there is no longer full risk-sharing conditional on a bad signal. Using Proposition 5, the expected utility of the protection buyer is given by

$$\begin{aligned} & \text{prob}[\bar{s}]u\left(E[\tilde{\theta}|\bar{s}] - \frac{\text{prob}[\underline{s}]A[\alpha^*(\underline{s})(R-C) + (1-\alpha^*(\underline{s}))\mathcal{P}]}{\text{prob}[\bar{s}]}\right) + \\ & \pi(1-\lambda)u\left(\bar{\theta} + \alpha^*(\underline{s})A - (1-\alpha^*(\underline{s}))A\frac{(1-\pi)R - \mathcal{P}}{\pi}\right) + (1-\pi)\lambda u(\underline{\theta} + \alpha^*(\underline{s})A + (1-\alpha^*(\underline{s}))AR) \end{aligned}$$

The derivative of the expected utility with respect to  $p$  is given by

$$\begin{aligned} & -\text{prob}[\bar{s}]u' \left( E[\tilde{\theta}|\bar{s}] - \frac{\text{prob}[\underline{s}]A[\alpha^*(\underline{s})(R-C) + (1-\alpha^*(\underline{s}))\mathcal{P}]}{\text{prob}[\bar{s}]} \right) \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A(1-\alpha^*(\underline{s}))\frac{\partial \mathcal{P}}{\partial p} + \\ & \frac{\pi(1-\lambda)}{\pi}u' \left( \bar{\theta} + \alpha^*(\underline{s})A - (1-\alpha^*(\underline{s}))A\frac{(1-\pi)R - \mathcal{P}}{\pi} \right) A(1-\alpha^*(\underline{s}))\frac{\partial \mathcal{P}}{\partial p} = \text{prob}[\underline{s}]A(1-\alpha^*(\underline{s}))\frac{\partial \mathcal{P}}{\partial p} \times \\ & \left[ u' \left( \bar{\theta} + \alpha^*(\underline{s})A - (1-\alpha^*(\underline{s}))A\frac{(1-\pi)R - \mathcal{P}}{\pi} \right) - u' \left( E[\tilde{\theta}|\bar{s}] - \frac{\text{prob}[\underline{s}]A[\alpha^*(\underline{s})(R-C) + (1-\alpha^*(\underline{s}))\mathcal{P}]}{\text{prob}[\bar{s}]} \right) \right] \end{aligned}$$

where we used  $\frac{\pi(1-\lambda)}{\pi} = \text{prob}[\underline{s}]$  and we again made use of the envelope theorem to claim  $\frac{\partial \alpha^*(\underline{s})}{\partial p} = 0$ . Since  $\alpha^*(\underline{s}) < 1$  (Proposition 6),  $1 - \alpha^*(\underline{s}) > 0$ . Using (50), and the fact that the limited liability constraints in states  $(\bar{\theta}, \bar{s})$  and  $(\bar{\theta}, \underline{s})$  are always slack (Proposition 2) and the incentive constraint after a bad signal binds (Proposition 4), we have that  $u'(\bar{\theta} + \tau(\bar{\theta}, \underline{s})) > u'(\bar{\theta} + \tau(\bar{\theta}, \bar{s}))$  or, equivalently, that the term in the square brackets above is positive. Since  $\frac{\partial \mathcal{P}}{\partial p} < 0$ , the expected utility under effort decreases in  $p$  when the limited liability constraint in state  $(\underline{\theta}, \underline{s})$  is binding.

We now show that the expected utility of the contract with risk-taking is increasing in  $p$ . Under risk-taking, the consumption of the protection buyer is equalized across all states. Therefore, using the optimal transfers from Proposition 10 in (27), the expected utility of the protection buyer under no effort is given by:

$$u \left( E[\tilde{\theta}] - \text{prob}[\underline{s}](1-p)A\mathcal{P} \right)$$

Using  $(1-p)A\mathcal{P} = R - C - pR$ , we have that the derivative of the expected utility with respect to  $p$  is given by

$$\text{prob}[\underline{s}]ARu' \left( E[\tilde{\theta}] - \text{prob}[\underline{s}](1-p)A\mathcal{P} \right) > 0$$

Lastly, note that as  $p \rightarrow \frac{R-C}{R}$  (or, equivalently, as  $\mathcal{P} \rightarrow 0$ ), the expected utility under risk-taking is strictly higher than the expected utility under effort. This is because the expected utility under risk-taking is approaching  $u \left( E[\tilde{\theta}] \right)$ , which is the first-best level of utility, while the expected utility under effort is strictly smaller than the first-best level of utility since  $A\mathcal{P} < (\pi - \underline{\pi})\Delta\theta$  and hence it is not possible to reach the first-best with effort after bad news (Lemma 1).

In sum, for  $p < \max\{\frac{R-C-1}{R-1}, \frac{1}{R}\}$ , the contract with effort is optimal. For  $p \rightarrow \frac{R-C}{R}$ , the contract with risk-taking is optimal. For  $\max\{\frac{R-C-1}{R-1}, \frac{1}{R}\} \leq p < \frac{R-C}{R}$ , the expected utility under effort is decreasing in  $p$  while the expected utility under risk-taking is increasing in  $p$ . Therefore, there exists a threshold value of  $p$ , denoted by  $\hat{p}$ , such that effort after bad news is optimal if and only if  $p \leq \hat{p}$ .

QED

### Proof of Proposition 12

With protection seller effort, there is full risk-sharing conditional on the realization of the signal  $\tilde{s}$  and we can write the objective function (6) as

$$U = \text{prob}[\bar{s}]u(E[\tilde{\theta} + \tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]) + \text{prob}[\underline{s}]u(E[\tilde{\theta} + \tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]). \quad (78)$$

Using the binding incentive and participation constraints, equations (15) and (16) express the expected transfer to protection buyers conditional on the signal,  $E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]$  and  $E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]$ , as a function of the margin  $\alpha(\underline{s})$  (recall that there is no margin call after a good signal). Writing the problem in terms of the expected transfers after a signal simplifies the exposition of the proof.

The first partial derivative of the objective function with respect to the margin is (for notational ease, we drop the reference to the  $\underline{s}$  in  $\alpha(\underline{s})$ ):

$$\frac{\partial U}{\partial \alpha} = \text{prob}[\bar{s}]\frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \alpha}\bar{u}' + \text{prob}[\underline{s}]\frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \alpha}\underline{u}', \quad (79)$$

where  $\underline{u}'$  and  $\bar{u}$  denote the marginal utility conditional on the bad and the good signal, respectively. The partial derivative of the expected transfer after a bad signal with respect to the margin is

$$\frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \alpha} = A[1 - \mathcal{P}(\alpha) + (1 - \alpha)\mathcal{P}'(\alpha)]. \quad (80)$$

When the derivative is positive, margins relax the incentive constraint. Define

$$X \equiv 1 - \mathcal{P}(\alpha) + (1 - \alpha)\mathcal{P}'(\alpha) \quad (81)$$

The derivative is positive if and only if  $X > 0$ . This is condition (41) in the text.

The partial derivative of the expected transfer after a good signal with respect to the margin is

$$\frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \alpha} = -\frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A[(R - C - 1) + X] \quad (82)$$

The derivative is negative when  $X > 0$  since  $R - C > 1$  (condition (2)). When  $X < 0$ , then the derivative may either be positive or negative, depending on how  $X$  compares to the opportunity cost of margins,  $R - C - 1$ .

Combining (80), (81) and (82), we can write (79) as

$$\frac{\partial U}{\partial \alpha} = \text{prob}[\underline{s}]A\bar{u}' \left[ \frac{\underline{u}'}{\bar{u}'} - \left( \frac{R - C - 1}{X} + 1 \right) \right] X$$

When  $X > 0$  then  $\frac{\partial U}{\partial \alpha} = 0$  yields the condition for an optimal interior margin in the proposition (when  $X < 0$  then  $\frac{\partial U}{\partial \alpha} < 0$  for sure since  $\frac{\underline{u}'}{\bar{u}'} \geq 1$ ). (Note that as in the linear cost case, it may be optimal not to use margins).

We now show that when  $\gamma < 0$ , then the optimization problem may not be well-behaved. The second partial derivative of the objective function (78) with respect to margins is

$$\begin{aligned} \frac{\partial^2 U}{\partial \alpha^2} = & \text{prob}[\bar{s}] \left[ \bar{u}'' \left( \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \alpha} \right)^2 + \bar{u}' \frac{\partial^2 E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \alpha^2} \right] \\ & + \text{prob}[\underline{s}] \left[ \underline{u}'' \left( \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \alpha} \right)^2 + \underline{u}' \frac{\partial^2 E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \alpha^2} \right] \end{aligned}$$

The first term in each squared bracket is negative (because of concave utility). A sufficient condition for a local maximum is therefore

$$\text{prob}[\bar{s}]\bar{u}' \frac{\partial^2 E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \alpha^2} + \text{prob}[\underline{s}]\underline{u}' \frac{\partial^2 E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \alpha^2} \leq 0$$

Using (80), (81) and (82) the condition becomes

$$\text{prob}[\underline{s}]A \frac{\partial X}{\partial \alpha} (\underline{u}' - \bar{u}') \leq 0.$$

Since  $\underline{u}' - \bar{u}' \geq 0$  (protections buyers may bear signal risk), the sufficient condition holds when  $\frac{\partial X}{\partial \alpha} \leq 0$  or, equivalently, when  $\gamma \geq 0$ . When  $\gamma < 0$  we cannot be sure that the first-order condition identifies a local maximum.

Finally, note that when  $\gamma \geq 0$  then  $1 > R - \frac{c}{1-p}$  is sufficient for  $X > 0$  for all  $\alpha$ .

QED

### Proof of Proposition 13

The first-order condition stipulates  $\frac{\partial U(\alpha^*, \gamma)}{\partial \alpha} = 0$  (for simplicity we consider only interior solutions,  $\alpha^* \in (0, 1)$ ). After total differentiation of this implicit function we obtain

$$\frac{d\alpha^*}{d\gamma} = -\frac{\frac{\partial^2 U}{\partial \alpha \partial \gamma}}{\frac{\partial^2 U}{\partial \alpha^2}}$$

When  $\alpha^*$  is a local maximum, then a more convex cost of effort leads to larger optimal margins,  $\frac{d\alpha^*}{d\gamma} > 0$ , if and only if  $\frac{\partial^2 U}{\partial \alpha \partial \gamma} > 0$ . This cross-partial derivative is

$$\begin{aligned} \frac{\partial^2 U}{\partial \alpha \partial \gamma} = & \text{prob}[\bar{s}] \left[ \bar{u}'' \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \gamma} \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \alpha} + \bar{u}' \frac{\partial^2 E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \alpha \partial \gamma} \right] \\ & + \text{prob}[\underline{s}] \left[ \underline{u}'' \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \gamma} \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \alpha} + \underline{u}' \frac{\partial^2 E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \alpha \partial \gamma} \right] \end{aligned}$$

Using (80), (81) and (82), the cross-partial derivative becomes

$$\begin{aligned} \frac{\partial^2 U}{\partial \alpha \partial \gamma} = & \text{prob}[\underline{s}] A \times \\ & \left[ -\bar{u}'' \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \gamma} [(R - C - 1) + X] + \underline{u}'' \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \gamma} X + \frac{\partial X}{\partial \gamma} (\underline{u}' - \bar{u}') \right] \end{aligned}$$

Moreover,

$$\begin{aligned} \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \gamma} &= \frac{\text{prob}[\underline{s}] (1 - \alpha)^2 A^2}{\text{prob}[\bar{s}] (1 - p)} > 0 \\ \frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \gamma} &= -\frac{(1 - \alpha)^2 A^2}{1 - p} < 0 \\ \frac{\partial X}{\partial \gamma} &= 2 \frac{(1 - \alpha) A}{1 - p} > 0 \end{aligned}$$

When  $\gamma \geq 0$  then  $\alpha^*$  is a local maximum and  $1 > R - \frac{c}{1-p}$  is sufficient for  $X > 0$ . And when  $X > 0$ , the cross-partial derivative is positive.

QED