Risk-sharing or risk-taking? Counterparty risk, incentives and margins*

Bruno Biais[†] Florian Heider[‡] Marie Hoerova[§]

This version: July 17, 2015

ABSTRACT

Derivatives activity, motivated by risk-sharing, can breed risk-taking. Bad news about the risk of the asset underlying the derivative increases the expected liability of a protection seller and undermines her risk-prevention incentives. This limits risk-sharing, and may create endogenous counterparty risk and contagion from news about the hedged risk to the balance sheet of protection sellers. 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.

*We would like to thank the Editor (Cam Harvey), the Acting Editor (Markus Brunnermeier), an Associate Editor and two anonymous referees, our discussants Ulf Axelson, Jonathan Berk, Sugato Bhattacharyya, Bruce Carlin, Simon Gervais, Artashes Karapetyan, Christine Parlour, Alp Simsek, Lauri Vilmi as well as numerous seminar and conference participants for their comments and suggestions. 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 (Grant 295484-TAP).

†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

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

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

JEL classification: G21, G22, D82

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

Derivatives activity has grown strongly over the past fifteen years. For example, credit default swaps (CDS), bilateral over-the-counter contracts used to insure credit risk, alone saw total notional amounts outstanding increase from around $180 billion in 1998 to a peak of over $60 trillion by mid-2008 (Acharya et al., 2012). But the insurance provided by derivatives is effective only if counterparties can honor their contractual obligations and do not default. When Lehman Brothers filed for bankruptcy protection in September 2008, it froze the positions of more than 900,000 derivative contracts (about 5% of all derivative transactions globally) in which Lehman Brothers was a party (Fleming and Sarkar, 2014).

What are the interactions between counterparty risk and derivatives activity? Can risk-sharing via derivatives perversely lead to risk-taking by financial institutions? How can derivatives activity be made more resilient to risk? In this paper, we explain how derivatives positions affect risk-taking incentives. We show how margin deposits and clearing arrangements can be designed to mitigate counterparty risk. We provide new empirical predictions about the extent of derivatives activity in a given financial environment and the default risk of institutions selling protection through derivatives.

Our model features risk-averse protection buyers who want to insure against a common exposure to risk (any idiosyncratic component of risk can be diversified among protection buyers themselves). To insure the common risk, they contact risk-neutral protection sellers whose assets can be risky, but who are not directly exposed to the risk buyers want to insure. Because of limited liability, protection sellers can make insurance payments only if their assets are sufficiently valuable. The value of a protection seller's assets is affected by her actions. Specifically, we assume protection sellers can prevent downside risk, and hence maintain a sufficient value for their assets, by exerting costly effort. For example, when choosing their investments they can carefully scrutinize their quality. Instead of careful and costly scrutiny, protection sellers can "shirk" and avoid the cost by relying on external, ready-made credit ratings or simple backward-looking measures of risk, as pointed out by Ellul and Yerramilli (2013). A failure of protection sellers to exert the risk-prevention effort (which we call "risk-taking") leads to counterparty risk for protection buyers.

Since financial institutions' balance sheets and activities are opaque and complex, lack of risk-prevention effort is difficult to observe and detect for outsiders. This creates a moral hazard problem for protection sellers, the key friction in our model.

Our model builds on two important characteristics of derivatives activity. First, during the life of a derivative contract, new information about the value of the underlying asset becomes available. Such news affect the expected pay-offs of the contracting parties: it makes the derivative position an asset for one party and a liability for the other. Second, derivative exposures, and hence the associated potential liabilities, can be large. According to the Quarterly Report on Bank Derivatives Activities by the Office of the Comptroller of the Currency, total credit exposure from derivatives reached more than \$1.5 trillion in 2008.[1] The total credit exposure of the top five financial institutions was two to ten times larger than their risk-based capital.

A key insight of our analysis is that a large derivative exposure undermines a protection seller's incentives to exert the risk-prevention effort when news makes the derivative position an expected liability for her. In that case, she bears the full cost of the risk-prevention effort while the benefit of this effort partly accrues to her counterparty in the form of payments from the derivative contract. This is reminiscent of debt-overhang (Myers, 1977) but there is an important difference. In our analysis, the liability arises endogenously in the context of an optimal contract and it only materializes when negative news occur.

The optimal contract takes one of two forms, depending on the severity of the moral hazard problem. Either the contract maintains protection sellers' risk-prevention incentives, but this comes at the cost of less ex-ante risk-sharing for protection buyers. Or it promises more risk-sharing but gives up on risk-prevention incentives, which creates counterparty risk for protection buyers. Thus, we show how the risk-sharing potential from derivatives contracts is limited either by the potential or by the actual presence of endogenous counterparty risk.[2]

The main focus of our paper is to characterize the optimal design of margin calls and central clearing, two institutional arrangements that aim to mitigate counterparty risk in derivatives. Both margins and central clearing received much attention in the regulatory overhaul of financial markets in the aftermath of the financial crisis. The Dodd-Frank Wall Street Reform Act in the U.S. and the European Market Infrastructure Regulation in Europe require certain derivative trades to occur via central clearing platforms (CCPs). There is, however, still considerable debate about the optimal design of CCPs for derivatives (see, e.g., Dudley, 2014, and Economist, 2014).

To examine the effects of central clearing, our model features a CCP that interposes between protection buyers and sellers. The benefit of the CCP is that it mutualizes the idiosyncratic part of counterparty risk. In a bilateral contract, each protection buyer is exposed to the counterparty risk of his own protection seller. The CCP instead pools the resources from all protection sellers. Any losses from the default of individual sellers are therefore shared across all protection buyers.

The CCP is also in charge of implementing margin calls. We emphasize the incentive role of margins. The party subject to a margin call has to deposit assets with the CCP. She no longer has control over the deposited assets, which are therefore "ring-fenced" from moral hazard. Risk-prevention effort only concerns the remaining, now smaller fraction of assets over which she still has control. The cost of risk-prevention effort is therefore lower, which improves risk-prevention incentives. While ring-fencing is the benefit of margins, it comes at a cost. The loss of control goes hand-in-hand with a loss of income. Safe assets on a margin account earn a lower return than risky assets left on financial institutions' balance sheets. Margins will therefore be used only when the ring-fencing benefit outweighs their cost, e.g., when the moral hazard problem is severe, or when the opportunity cost of depositing assets in the margin account is not too large.

Our analysis implies margins can be an attractive substitute to equity capital. 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. Consequently, margins allow protection sellers to engage in incentive-compatible derivative trading with less equity. An advantage of margins is their contingent nature. They are called only when individual derivative positions deteriorate.

Our mechanism design approach clarifies how two important reform proposals to make derivative markets more resilient, namely margins and central clearing, interact and need to be designed together. While central clearing allows mutualizing counterparty risk, margins provide incentives to avoid counterparty risk. Without margins, CCPs would bear too much risk and without a CCP, contracting parties would have to put up higher margins. And it is the CCP who must design and mandate the margin calls. Otherwise, there would be free-riding on the insurance it offers.

We also identify a channel through which derivatives activity can propagate risk. Without moral hazard, we assume for simplicity that the pay-offs from protection seller assets and protection buyer assets are independent. In contrast, with moral hazard, bad news about protection buyer assets can increase the likelihood of low pay-offs from protection seller assets, because bad news undermine protection sellers' risk-prevention incentives. Moral hazard in derivatives activity can therefore generate contagion (endogenous correlation) between two, otherwise unrelated, asset classes.

For example, prior to the recent crisis commercial banks frequently reduced their capital requirements by purchasing derivatives.[3] A bank exposed to sub-prime mortgages could purchase CDS on mortgages and save on regulatory capital. Conditional on the drop in real estate prices (which started well in advance of the crisis), those CDS contracts became expected liabilities for those institutions that sold them, typically investment banks. Our model predicts that financial institutions with larger short CDS positions exposed their balance sheets more to downside risks as bad news about the housing market emerged. This creates correlation between mortgage values and the values of financial institutions' assets without direct exposure to mortgage default. By contrast, those same institutions would not have increased their risk exposure after good news about the housing market. Importantly, in our model the exposure to downside risk is not the consequence of mistakes or incompetence. It is a calculated *choice* of trading-off ex-ante risk-sharing and downside risk exposure after bad news.

Our model generates new testable implications. First, we predict that derivatives contracts that offer ample insurance but increase exposure to counterparty risk are likely to be underwritten in a "benign" macroeconomic and financial environment. Second, the relation between derivatives exposures and the pledgeability of a financial institution's assets (measured, e.g., by the efficiency of its risk-management practices) is U-shaped. Financial institutions with an intermediate level of risk-management efficiency choose small derivatives exposures while financial institutions on the other two sides of efficiency spectrum choose large exposures. Third, optimal margins are higher when i) risk-free rates are high compared to the return on productive investment opportunities, and ii) risk-management costs increase strongly with the amount of assets under management.

While the financial insurance literature typically focuses on moral hazard on the part of the buyer of protection,[4] Thompson (2010) assumes moral hazard on the part of the seller of protection. Apart from that common assumption, our economic focus and modeling framework are very different from his. In Thompson (2010), i) the protection buyer is privately informed about his own risk and ii) the hidden action of the protection seller is the type of asset she invests in (liquid with low return, or illiquid but profitable in the long-term). In this context, moral hazard alleviates adverse selection and therefore *facilitates* the provision of insurance: High–risk protection buyers have incentives to reveal their type to induce the protection–seller to invest in the liquid asset, which is then available to pay the insurance when the loss occurs. In our set-up there is no adverse selection and moral hazard *hinders* the provision of insurance.

Other recent papers analyzing frictions in derivatives contracts include Bolton and Oehmke (2013) and Acharya and Bisin (2014). Bolton and Oehmke (2013) use our mechanism by which posting margins mitigates the moral hazard problem of the protection seller in a model which shows that effective seniority for derivatives transfers credit risk to the firm's debtholders that could be borne more efficiently by the derivative market. Acharya and Bisin (2014) analyze the externalities arising between several protection buyers when contracting with the same protection seller. They show how centralized clearing can internalize externalities among protection buyers, via optimally designed pricing schedules. This differs from our moral hazard setting where externalities are not a key issue, and quantities as well as prices must be controlled to restore incentives.

Our paper explains how derivatives activity, through its effect on incentives, can generate contagion between asset classes whose risk is independent in the absence of incentive problems. This novel form of contagion adds to the literature on shock propagation, which emphasizes interregional financial connections (Allen and Gale, 2000, Freixas, Parigi and Rochet, 2000), information contagion (Acharya and Yorulmazer, 2007, King and Wadhwani, 1990) and fire sales (Allen and Carletti, 2006, Cifuentes, Ferrucci and Shin, 2005).

Margins can be interpreted as a form of collateral. Collateral is usually analyzed in models in which agents borrow to finance investments (see, e.g, Bolton and Scharfstein (1990), Holmström and Tirole (1997), Acharya and Viswanathan (2011)). Our paper offers the first analysis of the incentive role of collateral in derivatives trading. This new context brings about new features that set margins apart from standard collateral. Standard collateral, say a house that backs up a mortgage, is transferred from the borrower to the creditor after decisions have been taken and pay-offs are realized, e.g., when the borrower defaults. By contrast, margin calls in our analysis, as in derivatives markets, occur before contracts mature, i.e., before final pay-offs are realized, and, importantly, before effort and risk-taking decisions are made.

Our modelling of moral hazard, where the agent chooses between effort and shirking is in line with Holmström and Tirole (1997, 1998), and we borrow from their analysis the terminology "pledgeable income," to refer to the future output that can be promised by the agent without jeopardizing her incentives. In our setting, however, incentives can be undermined by the arrival of information about the risk underlying the derivative contract before effort decisions are made, and this problem can be mitigated with margin calls. These key features of our model are absent from the standard moral hazard model studied in Holmström and Tirole (1997, 1998).

The remainder of the paper is organized as follows. The model is presented in Section 2, which also analyzes the benchmark case in which there is no moral hazard problem. We then analyze optimal contracting under moral hazard. Section 3 presents the solution when a buyer and a seller enter a bilateral contract without margins and without clearing. We use this set-up to show how moral hazard limits the provision of insurance through derivatives. Section 4 analyzes the optimal design of margins and central clearing, two distinct mechanisms to deal with counterparty risk in derivative contracts. Section 5 and 6 contains empirical and policy implications. Section 7 concludes. Proofs as well as extensions and robustness are in the Online Appendix.

I. Model and First-Best Benchmark

A. 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 $t = 2$.[5] For simplicity, we assume $\tilde{\theta}$ can only take on two values: $\bar{\theta}$ with probability π 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.[6]

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 such risk-prevention effort, the per unit return $\tilde{R}_j$ is R with probability one. If she does not exert the risk-prevention 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. Combined with limited liability, effort unobservability generates moral hazard.

When exerting risk-prevention effort, the agent incurs nonpecuniary costs C per unit of assets under management at $t = 1$.[7] Because protection seller assets are riskier without costly effort, we also refer to the decision not to exert effort as "risk-taking".[8] Undertaking effort is efficient,

$$R - C > pR, \tag{1}$$

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

$$R - C > 1. \tag{2}$$

For simplicity, conditional on effort, $\tilde{R}_j$ is independent across sellers and independent of protection buyers' risk $\tilde{\theta}$. To allow protection sellers that exert effort to fully insure buyers, we assume

$$AR > \pi\Delta\theta. \tag{3}$$

Advance information. At the beginning of $t = 1$, before investment and effort 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, $\tilde{s}$ can be the real-estate price index. Denote the conditional probability of a correct signal by

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

Hence, $\text{prob}[\bar{s}] = \lambda\pi + (1-\lambda)(1-\pi)$ and $\text{prob}[\underline{s}] = (1-\lambda)\pi + \lambda(1-\pi)$.

The probability π of a good outcome $\bar{\theta}$ for protection buyers' 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,

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

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 $\lambda = 1$, the signal is perfectly informative.

Contracts, margins and central clearing. The contract specifies separate transfers to protections buyers, τ^B, and to protection sellers, τ^S. Positive transfers $\tau^S, \tau^B > 0$ represent payments made to the seller and to the buyer, while negative transfers represent payments made by the seller and by the buyer. In the case of contracting without a CCP, there is a single contract between a protection buyer and a protection seller so that $\tau^S = -\tau^B$. In the case of contracting with a CCP, the CCP interposes between contracting parties. Thus, the contract between a protection buyer and a protection seller is transformed into two contracts (a process called novation). One contract is between the seller and the CCP specifying a transfer τ^S. The other contract is between the buyer and the CCP specifying a transfer τ^B.

The transfers τ^S and τ^B are agreed upon at $t = 0$, they occur at $t = 2$, and are contingent on all available information at that time. This information consists of the buyers' risk $\tilde{\theta}$, the signal $\tilde{s}$ and the return on the protection seller's assets $\tilde{R}_j$. Hence, we write $\tau^S(\tilde{\theta}, \tilde{s}, \tilde{R}_j)$ and $\tau^B(\tilde{\theta}, \tilde{s}, \tilde{R}_j)$. 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), they can be interpreted as transfers specified by derivative contracts.

The contract can also specify margin deposits. Margin deposits are implemented as escrow accounts set up by a protection buyer or by a CCP. Importantly, we assume that margin deposits are observable and contractible, and that contractual provisions calling for margin deposits are enforceable.[9] Because a protection buyer or a CCP have no ability to manage risky, opaque assets, margin deposits must be satisfied with safe, transparent assets (e.g., cash or Treasuries, which are not subject to information asymmetry problems).[10] One can therefore interpret margins as an institutional arrangement that affects the time-1 split of a seller's balance sheet between transparent assets and opaque investments. Margins "ring-fence" a fraction of 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 $t = 1$, i.e., the signal $\tilde{s}$.[11] We denote the fraction of a protection seller's balance sheet deposited on the margin account by $\alpha(s)$ so that

$$\alpha(s) \in [0, 1] \quad \forall s. \tag{4}$$

Transfers from a protection seller are constrained by limited liability. Each protection seller j 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,

$$-\tau^S(\theta, s, R_j) \leq \alpha(s)A + (1 - \alpha(s))AR_j, \quad \forall(\theta, s, R_j, j). \tag{5}$$

In case of contracting with a CCP, because the novation creates separate contracts for protection buyers and sellers with the CCP, there can be mutualization. That is, the transfers to protection buyers depend not only on a protection seller's individual asset return $\tilde{R}_j$, as is the case in a bilateral contract without the CCP, but they depend on all sellers' asset returns $\tilde{R}$ because they affect the amount of resources available to the CCP to distribute among its members. The CCP is still subject to budget-balance (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):

$$\tau^B(\theta, s, R) \leq -\tau^S(\theta, s, R), \quad \forall(\theta, s, R). \tag{6}$$

In the case of contracting with a CCP we adopt 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. In the case of contracting without a CCP each individual protection buyer maximizes his expected utility subject to the participation constraint of his protection seller.[12]

The sequence of events is summarized in Figure 1.

Insert Figure 1 here

B. First-best: No moral hazard

In this subsection we consider the case in which protection sellers' risk-prevention 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 that arise when protection seller's risk-prevention effort is not observable.

In the first-best protection sellers are requested to exert risk-prevention effort when offering protection since doing so increases the resources available for risk-sharing (see (1)). Margins are not used because they are costly (see (2)) and offer no benefit when risk-prevention effort is observable. Since protection sellers exert effort, their assets generate return $\tilde{R}_j = R$ for sure and there is no counterparty risk. Therefore there is no benefit to the mutualization of risk and the optimal contract designed with or without a CCP is the same. Since $\tilde{R}_j = R$ for sure, we drop the reference to R in the transfers for ease of notation. The contract specifies transfers to buyers and sellers, $\tau^B(\tilde{\theta}, \tilde{s})$ and $\tau^S(\tilde{\theta}, \tilde{s})$, to maximize buyers' utility

$$E[u(\tilde{\theta} + \tau^B(\tilde{\theta}, \tilde{s})] \quad (7)$$

subject to the limited liability (5) and feasibility (6) constraints, as well as the constraint that protection sellers enter the contract. By entering (and exerting effort) sellers obtain $E[\tau^S(\tilde{\theta}, \tilde{s})] + A(R - C)$. If they do not enter the contract and sell protection, they obtain $A(R-C)$.[13] Therefore a protection seller's participation constraint under effort is

$$E[\tau^S(\tilde{\theta},\tilde{s})] \geq 0. \tag{8}$$

Proposition 1 states the first-best outcome.

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*

$$\begin{aligned}
\tau^B(\bar{\theta},\bar{s}) &= \tau^B(\bar{\theta},\underline{s}) = E[\tilde{\theta}] - \bar{\theta} = -(1-\pi)\,\Delta\theta < 0 \\
\tau^B(\underline{\theta},\bar{s}) &= \tau^B(\underline{\theta},\underline{s}) = E[\tilde{\theta}] - \underline{\theta} = \pi\Delta\theta > 0 \\
\tau^B(\theta,s) &= -\tau^S(\theta,s) \quad \forall(\theta,s)
\end{aligned}$$

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 return θ 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, $E[\tilde{\theta}]$. Each transfer to a protection buyer is matched by an opposite transfer from a protection seller. The first-best insurance contract is actuarially fair since the expected transfer from protection sellers to protection buyers is zero, $E[\tau^B(\tilde{\theta},\tilde{s})] = -E[\tau^S(\tilde{\theta},\tilde{s})] = 0$. By our assumption (3), the resources generated by the protection sellers are large enough to fully insure protection buyers.

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

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 at $t=1$ in which, before the realization of the signal, a protection seller uses some of her initial assets A to acquire a protection buyer's asset at price S. Because there is no discounting, this is equivalent, from the point of view of 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 θ 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.

II. Bilateral contracts without margins

In the previous section we examined the hypothetical case in which protection sellers' risk-prevention effort is observable and can therefore be requested by protection buyers. We now move on to the more realistic situation in which risk-prevention effort is not observable. We first analyze the case of a bilateral contract between a buyer and a seller, without the presence of a CCP and without margins. We use this set-up to illustrate the inefficiencies in risk-sharing introduced by the moral-hazard in protection sellers' risk-management. In the next section, we show how contracting with a CCP and the availability of margin deposits address those inefficiencies.

If a protection buyer wants a protection seller to exert risk-prevention effort when it is unobservable, then it must be in the seller's own interest to exert effort after seeing the signal s about the buyer's risk $\tilde{\theta}$. The incentive compatibility constraint under which a protection seller exerts effort after observing s is:

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

The left-hand side is a protection seller's expected payoff if she exerts risk-prevention effort. The effort costs C per unit of assets under management, 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. Here we use the index j for the random return of a protection seller's assets because each protection buyer enters a bilateral contract with a single protection seller j.

Without effort, her assets under management return R with probability p and zero with probability $1-p$. In order to relax the incentive constraint as much as possible, the buyer does not pay the seller when $\tilde{R}_j = 0$: $\tau^S(\tilde{\theta}, \tilde{s}, 0) = 0$. With effort, protection seller assets are safe, with $\tilde{R}_j = R$. For brevity, we write $\tau^S(\tilde{\theta}, \tilde{s})$ as $\tau^S(\tilde{\theta}, \tilde{s})$. Then, the incentive constraint after observing s is

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

or, using the notion of "pledgeable return" $\mathcal{P}$ (see Holmström and Tirole, 1997),

$$\mathcal{P} \equiv R - \frac{C}{1-p}, \tag{9}$$

the incentive compatibility constraint rewrites as

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

The right-hand side is what the protection seller expects to pay the protection buyer after seeing the realization of the signal s of the buyer's risky asset $\tilde{\theta}$. The right-hand side is positive when conditional on the signal, the protection seller expects on average to make transfers to the protection buyer. If after seeing the signal she expects on average to receive transfers from the protection buyer, then the right-hand side is negative and the incentive constraint does not bind. This is an important observation to which we return later. The left-hand side is the amount that a protection seller can pay (or pledge) to the protection buyer without undermining her incentive to exert risk-prevention effort. 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 left-hand side is positive because the assumption that effort is efficient, condition (1), ensures positive pledgeable return, $\mathcal{P} > 0$.

When the pledgeable return $\mathcal{P}$ is sufficiently high, a protection seller's 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-prevention effort is not observable, the first-best can be achieved if and only if the pledgeable return on assets is high enough:*

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

The threshold for the pledgeable return on assets, beyond which full risk-sharing is possible despite protection seller moral hazard, 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 λ 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 the first-best is not attainable and, moreover, the signal is sufficiently informative. In particular, we assume that:

$$\lambda \geq \lambda^* \equiv \frac{1 - \sqrt{p}}{1 - p} > \frac{1}{2}. \tag{12}$$

While relatively mild,[14] this assumption simplifies the analysis by focusing on the case in which the moral hazard problem is relatively severe.

A. Risk-prevention effort after both signals

In this section, we study the contract that provides incentives to a protection seller to exert risk-prevention effort after both a good and a bad signal. The contract specifies transfers between a protection buyer and a protection seller, $\tau^S(\tilde{\theta}, \tilde{s}) = -\tau^B(\tilde{\theta}, \tilde{s})$, that maximize a buyer's utility (7) subject to the limited liability (5), participation (8) and incentive (10) constraints.

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

$$AR + \frac{\text{prob}[\underline{s}]\,(1-p)\,A\mathcal{P}}{1-\text{prob}[\underline{s}]\,(1-\underline{\pi})\,(1-p)} > \bar{\pi}\Delta\theta. \tag{13}$$

The assumption guarantees, as we will show, a slack limited liability constraint for transfers from a protection seller to a protection buyer when there is a good signal, $\bar{s}$, but the buyer's asset return is low, $\underline{\theta}$. This assumption is satisfied, for example, whenever a seller's assets under effort return enough relative to the risk exposure being hedged, $AR > \Delta\theta$.

The next proposition states that the moral hazard problem only matters after a bad signal, and that the seller breaks even on the contract.

PROPOSITION 2: *In the optimal contract with risk-prevention effort, the incentive constraint (10) binds after a bad signal, but is slack after a good signal. The participation constraint (8) binds.*

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-prevention effort. She has to bear the full cost of effort while the benefit of staying solvent accrues in part to the protection buyer in the form of the (likely) transfer to him. This is similar to the debt-overhang effect (Myers, 1977).[15]

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-prevention effort.

The binding participation constraint (8) makes the contract actuarilly fair. The sum of all transfer payments is zero.

The next proposition characterizes the optimal transfers in each state for a bilateral contract without margins.

PROPOSITION 3: *In the optimal contract with risk-prevention effort, the transfers* *to a protection buyer after a good signal are*

$$\tau^B(\bar{\theta}, \bar{s}) = (E[\tilde{\theta}|\bar{s}] - \bar{\theta}) - \frac{prob[\underline{s}]}{prob[\bar{s}]} A\mathcal{P} < 0, \qquad (14)$$
$$\tau^B(\underline{\theta}, \bar{s}) = (E[\tilde{\theta}|\bar{s}] - \underline{\theta}) - \frac{prob[\underline{s}]}{prob[\bar{s}]} A\mathcal{P} > 0,$$

and after a bad signal, they are

$$\tau^B(\bar{\theta}, \underline{s}) = (E[\tilde{\theta}|\underline{s}] - \bar{\theta}) + A\mathcal{P} < 0 \qquad (15)$$
$$\tau^B(\underline{\theta}, \underline{s}) = (E[\tilde{\theta}|\underline{s}] - \underline{\theta}) + A\mathcal{P} > 0.$$

Under condition (13), all limited liability constraints are slack.

In the optimal contract with effort, 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. However, in contrast with the first-best, transfers now vary with the signal. This is because it is more difficult to provide incentives to a protection seller after a bad signal. Thus, incentive compatibility reduces the transfers that can be requested from a protection seller. Correspondingly, protection buyers are exposed to signal risk. Their consumption is larger after a good signal than after a bad signal.

Cross-subsidization across signals mitigates the impact of signal risk, but only imperfectly because of 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. The cross-subsidization is given by the second component of each transfer in (14) and (15) (the terms not in round brackets). If the contract was written after that information is observed, such cross-subsidization would be not be possible. This would reduce the scope for insurance, in line with the Hirshleifer (1971) effect.

In the first-best, the transfers depend only on the realization of θ and the optimal contract can be implemented with a simple forward contract. In contrast, with moral hazard and risk-prevention effort after both signals, the transfers depend on the realizations of θ *and* s. The optimal contract can be implemented by the sale of a forward contract on the underlying asset θ 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 light of the liability from the forward contract on θ.[16]

B. No risk-prevention effort after a bad signal (risk-taking)

Maintaining incentive compatibility after a bad signal reduces risk-sharing. A protection buyer may find this reduction in insurance too costly. He may instead choose to accept shirking on risk-prevention effort (risk-taking) by the protection seller in exchange for a better sharing of the risk associated with $\tilde{\theta}$. In this section, we characterize the optimal contract a seller does not exert risk-prevention effort after a bad signal.

As before, a protection seller's incentives to exert effort are intact after a good signal so that $\tilde{R}_j = R$ for all j. After a bad signal, the seller now shirks so that $\tilde{R}_j = R$ with probability p and $\tilde{R}_j = 0$ with probability $1 - p$. Hence, the transfers $\tau^B = -\tau^S$ now must be contingent on the realization of $\tilde{R}_j$. The objective function of a protection buyer is given by:

$$\begin{aligned}\max_{\tau}\ & \pi\lambda u(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}, R)) + (1-\pi)(1-\lambda)u(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}, R)) \\ & + \pi(1-\lambda)[pu(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, R)) + (1-p)u(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, 0))] \\ & + (1-\pi)\lambda[pu(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s}, R)) + (1-p)u(\underline{\theta})]\end{aligned} \tag{16}$$

After a bad signal, a seller's assets return zero with probability $1 - p$ and she cannot make any transfers to a protection buyer. In that case, when the bad state $\underline{\theta}$ is realized, a protection buyer would like to receive an insurance payment, but a protection seller is unable to pay and so the transfer for this state is optimally set to zero, $\tau^B(\underline{\theta}, \underline{s}, 0) = -\tau^S(\underline{\theta}, \underline{s}, 0) = 0$. The buyer is uninsured and consumes $\underline{\theta}$. It may, however, be optimal for a buyer to make a transfer to his seller when $R_j = 0$

and the good state $\bar{\theta}$ is realized, i.e., it may be optimal to set $\tau^B(\bar{\theta}, \underline{s}, 0) < 0$.

A seller's incentive constraint after a good signal is unchanged (condition (10) with $\tilde{s} = \bar{s}$). After a bad signal, however, she must now prefer not to exert effort:

$$E[\tau^S(\tilde{\theta}, \underline{s}, R)] + A\,(R - C) \leq pE[\tau^S(\tilde{\theta}, \underline{s}, R) + AR] + (1-p)E[\tau^S(\tilde{\theta}, \underline{s}, 0)],$$

or, equivalently,

$$A\mathcal{P} \leq -E[\tau^S(\tilde{\theta}, \underline{s}, R)] + E[\tau^S(\tilde{\theta}, \underline{s}, 0)]. \tag{17}$$

A seller's participation constraint with risk-taking is

$$\text{prob}[\underline{s}](1-p)A\mathcal{P} \leq \text{prob}[\bar{s}]E[\tau^S(\tilde{\theta}, \bar{s}, R)] + \text{prob}[\underline{s}](pE[\tau^S(\tilde{\theta}, \underline{s}, R)] + (1-p)E[\tau^S(\tilde{\theta}, \underline{s}, 0)]). \tag{18}$$

The expected transfer to a seller (right-hand side) is positive and hence, the contract with no effort after a bad signal is actuarially unfair. If a seller enters the position, she must be compensated for the potential efficiency loss due to the lack of effort after bad news (left-hand side). The higher the pledgeable income, the greater is the efficiency loss generated by risk-taking after a bad signal and the more actuarially unfair is the contract.

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

PROPOSITION 4: *If risk-taking (no effort) is preferred to effort after a bad signal, then the optimal contract provides full insurance except when a seller defaults in the $\underline{\theta}$ state. The transfers are given by $\tau^B(\underline{\theta}, \underline{s}, 0) = 0$ and*

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

Under condition (13), all limited liability constraints are slack.

Except when a protection seller defaults and the value of a buyer's assets is low $\underline{\theta}$, the consumption of the buyer is equalized across states. But when a protection seller defaults and $\underline{\theta}$ occurs, her protection buyer cannot receive any insurance payment and is therefore exposed to counterparty risk. The contract is actuarially unfair but there are no rents to a protection seller since the participation constraint binds. A protection buyer receives a transfer from his seller in the bad state $\tilde{\theta} = \underline{\theta}$ (if she does not default) and makes a transfer to her in the good state $\tilde{\theta} = \bar{\theta}$: $\tau^B(\underline{\theta}, \tilde{s}, R) > 0 > \tau^B(\bar{\theta}, \tilde{s}, \tilde{R})$.

Unlike in the contract with risk-prevention effort after a bad signal, the contract without such effort does not react to the signal, i.e., $\tau(\tilde{\theta}, \bar{s}) = \tau(\tilde{\theta}, \underline{s})$. It can therefore be implemented with a single forward (as in the first-best). The protection buyers must, however, sell the forward at a discount relative to the expected value of the underlying risk in order to compensate the protection seller for the loss of income in case of default.

Finally, whether the contract without risk-prevention effort after the bad signal is preferred to the contract with such effort after both signals depends on parameters. We defer the characterization of the choice between these two types of contracts to the next section, after we have derived the second-best contracts with margins and clearing.

III. Centrally cleared contracts with margins

In this section, we analyze the optimal design of contracts with margins and central clearing. These are two distinct mechanisms to deal with counterparty risk in derivatives.

Margins ring-fence a protection seller's assets from moral hazard. Such ring-fencing can serve two purposes, depending on whether the optimal contract leads to risk-prevention effort after both signals, or whether it leads to no such effort after a bad signal. In the contract with effort after both signals, margins can improve risk-prevention incentives ex ante and enhance the scope for risk-sharing. In the contract with no effort after a bad signal, margins can provide resources ex post in case of a seller's default.

Central clearing allows mutualization of idiosyncratic counterparty risk. Since the CCP interposes itself between a protection buyer and a protection seller, a protection buyer is no longer exposed to the default risk of her own protection seller. Even if some protection sellers default, the CCP can pool resources from those protection sellers who do not default and still make transfers to all protection buyers, thus providing insurance against counterparty risk.

We show that margins and clearing interact and solve for the optimal combination of the two. In the contract with effort, margins are used to improve ex ante risk-prevention incentives. Clearing is superfluous as there is no counterparty risk on the equilibrium path in our simple set-up. In the contract with no effort after a bad signal, clearing mutualizes counterparty risk. Margins are not used because risk-prevention incentives are ignored after a bad signal. It is more efficient to generate the resources for insurance through mutualization (all counterparty risk is idiosyncratic) rather than through margin calls.

A. Margins and clearing with effort after both signals

In this section, we study the contract with margins and clearing that provides incentives to protection sellers to exert risk-prevention effort both after a good and a bad signal. A margin call, requesting that a fraction $\alpha(\tilde{s})$ of a protection seller's balance sheet be deposited on the margin account, can be made at $t = 1$, after the signal $\tilde{s}$ is observed. A centrally cleared contract is conditional on all the sellers' asset returns $\tilde{R}$, unlike a bilateral contract which is conditional on an individual return, $\tilde{R}_j$.

With margins and clearing, therefore, the participation and incentive constraints are modified as follows. A protection seller's participation constraint is

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

Because protection sellers exert effort on the equilibrium path, 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, we have

$$E[\tau^S(\tilde{\theta}, \tilde{s})] \geq E[\alpha(\tilde{s})]A(R - C - 1), \tag{19}$$

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.

A seller's incentive constraint after observing s is now given by

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

or, after re-arranging,

$$\alpha(s)A + (1 - \alpha(s))A\mathcal{P} \geq E[-\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = s]. \tag{20}$$

The right-hand side is what protection sellers expect to pay to the CCP after seeing the signal about buyers' risk.[17] The left-hand side is the sum of the pledgeable return on i) the assets deposited on the margin account and on ii) those left under protection sellers' management. The pledgeable return on assets deposited on the margin account is equal to their physical return since they are "ring-fenced" from moral hazard in risk-management.

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 (7) subject to the feasibility constraints (4) and (6), limited liability constraints (5), and the participation (19) and incentive (20) constraints.

In what follows, we characterize the optimal contract with risk-prevention effort after both signals.

PROPOSITION 5: *In the optimal contract with risk-prevention effort after both signals, the feasibility constraints (6) bind for all (θ, s) and the participation constraint (19) binds. Margins are not used after a given signal s if the incentive constraint for* *that signal is slack or if the moral hazard is not severe to begin with, i.e.,* $\mathcal{P} \geq 1$.

The CCP passes on all resources available for insurance to protection buyers, $\tau_B = -\tau_S$, and protection sellers break even.

When the moral hazard is severe, $\mathcal{P} < 1$, then depositing assets on the margin account relaxes the incentive constraint (20) and thus allows for higher transfers to protection buyers. This is the benefit of margins. But assets deposited on the margin account are costly as they incur an opportunity cost $R - C - 1$ to protection sellers. When the incentive constraint after a given signal s is slack, then margins are not used because depositing assets on the margin account offers no incentive benefit and only incurs the opportunity cost. Similarly, margins are not used when the pledgeable return of assets under management is higher than the pledgeable return of assets deposited on the margin account, $\mathcal{P} \geq 1$.

The next proposition characterizes optimal transfers between the CCP, protection buyers and protections sellers that always exert risk-prevention effort. As in the contract without margins and clearing (Section II.A), the incentive constraint after a good signal is slack in the optimum. By Proposition 5, there is no margin call after a good signal, $\alpha(\bar{s}) = 0$. We now state optimal transfers as a function of a margin call after a bad signal, $\alpha(\underline{s})$, and examine the determination of the optimal margin thereafter.

PROPOSITION 6: *In the optimal contract with risk-prevention effort after both signals, the incentive constraint after a bad signal (20) binds. The transfers to protection buyers after a good signal are*

$$\tau^B(\bar{\theta}, \bar{s}) = (E[\tilde{\theta}|\bar{s}] - \bar{\theta}) - \frac{prob[\underline{s}]}{prob[\bar{s}]} A \left[\alpha(\underline{s})\,(R - C) + (1 - \alpha(\underline{s}))\mathcal{P}\right] < 0, \tag{21}$$
$$\tau^B(\underline{\theta}, \bar{s}) = (E[\tilde{\theta}|\bar{s}] - \underline{\theta}) - \frac{prob[\underline{s}]}{prob[\bar{s}]} A \left[\alpha(\underline{s})\,(R - C) + (1 - \alpha(\underline{s}))\mathcal{P}\right] > 0.$$

If the limited liability constraint is slack in state $(\underline{\theta}, \underline{s})$*, the transfers to protection* *buyers after a bad signal are*

$$\tau^B(\bar{\theta},\underline{s}) = (E[\tilde{\theta}|\underline{s}] - \bar{\theta}) + A\left[\alpha(\underline{s}) + (1-\alpha(\underline{s}))\mathcal{P}\right] < 0 \tag{22}$$
$$\tau^B(\underline{\theta},\underline{s}) = (E[\tilde{\theta}|\underline{s}] - \underline{\theta}) + A\left[\alpha(\underline{s}) + (1-\alpha(\underline{s}))\mathcal{P}\right] > 0.$$

Otherwise, the transfers after a bad signal are

$$\tau^B(\bar{\theta},\underline{s}) = \alpha(\underline{s})A - (1-\alpha(\underline{s}))A\frac{(1-\underline{\pi})R - \mathcal{P}}{\underline{\pi}} \tag{23}$$
$$\tau^B(\underline{\theta},\underline{s}) = \alpha(\underline{s})A + (1-\alpha(\underline{s}))AR > 0.$$

If the limited liability constraint in state $(\underline{\theta},\underline{s})$ is slack, then there is full risk-sharing given the signal, as in Section II.A. As before, the binding incentive constraint after a bad signal introduces risk across signals: protection buyers' consumption is larger after a good signal than after a bad signal. Unlike in Section II.A, the insurance offered is not actuarially fair. It involves a premium, to compensate protection sellers for the efficiency loss induced by margins: prob$[\underline{s}]\alpha(\underline{s})A(R-C-1)$. This premium is equal to the expectation of the second component of the transfers in (21) and (22).

The structure of the transfers in (23) is different. When limited liability binds in state $(\underline{\theta},\underline{s})$, full risk-sharing conditional on the signal is no longer possible because protection sellers' resources in state $(\underline{\theta},\underline{s})$ are insufficient. Conditional on a bad signal, the transfers in (23) implement whatever risk-sharing is still possible given the binding limited liability constraint.

To analyze the optimal amount of margin 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 φ, we have

$$\varphi \equiv \frac{u'(\bar{\theta} + \tau^B(\bar{\theta},\underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta},\bar{s}))} \tag{24}$$

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

Given the transfers in Proposition 6, φ is a known function of exogenous variables and $\alpha(\underline{s})$: the denominator of φ is increasing in $\alpha(\underline{s})$ and the numerator is decreasing in $\alpha(\underline{s})$. Hence, φ is decreasing in $\alpha(\underline{s})$. Higher margins reduce φ, as they reduce the wedge between consumption after a good and a bad signal, i.e., they improve insurance against signal risk. Optimal margins balance the insurance against signal risk against the benefit and cost of using margins.

PROPOSITION 7: *If $\mathcal{P} > 1$ or $\varphi(0) < 1 + \frac{R-C-1}{1-\mathcal{P}}$, margins are not used. Otherwise, there are two cases. If*

$$\varphi\left(1 - \frac{\underline{\pi}\Delta\theta}{A\,(R-\mathcal{P})}\right) < 1 + \frac{R-C-1}{1-\mathcal{P}}, \tag{25}$$

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

$$\varphi(\alpha^*(\underline{s})) = 1 + \frac{R-C-1}{1-\mathcal{P}}, \tag{26}$$

while, if (25) does not hold, the optimal margin solves

$$\varphi(\alpha^*(\underline{s})) = 1 + \frac{R-C-1}{1-\mathcal{P}} + \frac{1-\underline{\pi}}{1-\mathcal{P}} \frac{u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) - u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))}. \tag{27}$$

Lastly, putting all assets of a protection seller in the margin account cannot be optimal:

$$\alpha^*(\underline{s}) < 1. \tag{28}$$

The right-hand side of (26) reflects the trade-off between the cost and benefit of margins. The numerator, $R - C - 1$, is the opportunity cost of depositing a margin. The denominator, $1 - \mathcal{P}$, is the incentive benefit of ring-fencing assets from moral hazard.

Figure 2 illustrates how margins are optimally set (condition (26)). The figure is useful for examining graphically, for example, the effect of an increase in p, which reduces pledgeable income $\mathcal{P}$. The decrease in $\mathcal{P}$ shifts the curve φ upwards while shifting $1 + \frac{R-C-1}{1-\mathcal{P}}$ downwards. This raises the optimal margin in (26). When incentive problems become more severe, margins are needed more to relax the incentive constraint.

Insert Figure 2 here

When margins are as in (26), consistency requires that there be enough resources to provide full insurance conditional on the signal. This is the case if (25) holds. Consistent with intuition, (25) holds when R is large. When there is full risk-sharing conditional on the signal, the last term on the right hand-side of (27) is 0. In that case, (27) simplifies to (26).

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 (27) is strictly positive and margins are lower than when the limited liability condition is slack because (taking effort as given) margins reduce the amount of resources 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 intuition for why $\alpha^*(\underline{s}) < 1$ is as follows. When assets are put in the margin account, they earn a lower return than under the management of a protection seller exerting effort. This reduces the resources available to pay insurance to a protection buyer. When $\alpha^*(\underline{s}) = 1$ there is a severe dearth of resources to pay insurance. In fact, the entire pool of assets in the margin account will be used to pay insurance to protection buyers when $\underline{\theta}$ and hence, limited liability binds. In this case, as can be seen by inspecting (23) for $\alpha^*(\underline{s}) = 1$, the transfers are highly constrained. In fact, they are so constrained that very little risk-sharing can be achieved and a contract with $\alpha^*(\underline{s}) = 1$ is suboptimal.

B. Margins and clearing with no effort after a bad signal

Implementing risk-prevention effort necessitates a reduction in risk-sharing. Protection buyers may therefore instead choose to accept no effort after a bad signal (risk-taking) by protection sellers in exchange for a better sharing of the risk associated with $\tilde{\theta}$. In this section, we characterize the optimal risk-taking contract with margins and clearing.

After a good signal, protection sellers exert risk-prevention effort so that $\tilde{R}_j = R$ for all j. After a bad signal, protection sellers do not not exert risk-prevention effort so that $\tilde{R}_j = R$ with probability p and $\tilde{R}_j = 0$ with probability $1 - p$. Hence, the transfer τ^S between a protection seller and the CCP now is contingent on the realization of $\tilde{R}_j$. But in contrast to a bilateral contract (Section II.B), the transfer τ^B between a protection buyer and the CCP 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 depending on whether a good or a bad signal occurs. After a bad signal, only a proportion p of protection sellers generate a 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 $\tau^S(\tilde{\theta}, \tilde{s}, \tilde{R}_j)$, to maximize buyers' utility

$$\begin{aligned} &\pi\lambda u(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}, R)) + (1-\pi)(1-\lambda)u(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}, R)) \\ &+ \pi(1-\lambda)u(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, pR)) + (1-\pi)\lambda u(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s}, pR)) \end{aligned} \tag{29}$$

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

With risk-taking and mutualization, the feasibility constraints of the CCP after a good and a bad signal are given by

$$\tau^B(\theta, \bar{s}, R) \leq -\tau^S(\theta, \bar{s}, R) \quad \forall(\theta, \bar{s}) \tag{30}$$

$$\tau^B(\theta, \underline{s}, pR) \leq -p\tau^S(\theta, \underline{s}, R) - (1-p)\tau^S(\theta, \underline{s}, 0) \quad \forall(\theta, \underline{s}). \tag{31}$$

The seller's incentive constraint after a good signal is as before (condition (20) with $\tilde{s} = \bar{s}$), whereas after a bad signal, the seller must prefer not to exert effort

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

or, equivalently,

$$(1 - \alpha(\underline{s}))A\mathcal{P} \leq -E[\tau^S(\tilde{\theta}, \underline{s}, R)] + E[\tau^S(\tilde{\theta}, \underline{s}, 0)]. \tag{32}$$

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

$$\begin{aligned} &\text{prob}[\bar{s}]\alpha(\bar{s})A(R - C - 1) + \text{prob}[\underline{s}]\alpha(\underline{s})A(pR - 1) + \text{prob}[\underline{s}](1 - p)A\mathcal{P} \\ &\leq \text{prob}[\bar{s}]E[\tau^S(\tilde{\theta}, \bar{s}, R)] + \text{prob}[\underline{s}](pE[\tau^S(\tilde{\theta}, \underline{s}, R)] + (1 - p)E[\tau^S(\tilde{\theta}, \underline{s}, 0)]) \end{aligned} \tag{33}$$

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 arises because of two factors: 1) costly margins after a 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 $\text{prob}[\underline{s}](1 - p)$. As before, the contract with no effort after a bad signal is actuarially unfair.

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. It also states that protection sellers earn no rents and all resources available for insurance from protection sellers are passed on to protection buyers.

PROPOSITION 8: *In the optimal contract with no effort after a bad signal (risk-taking), the feasibility constraints bind for all (θ, s) and the participation constraint binds. 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 a 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 protection seller assets on the margin account, $\alpha = 1$. But protection buyers can do at least as well by inducing effort after a bad signal since, there too, $\alpha = 1$ can be selected (but, as we know from Proposition 7, 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 no effort after a bad signal.

PROPOSITION 9: *If $pR < 1$, then no effort after a bad signal (risk-taking) is sub-optimal. Otherwise, the optimal contract with risk-taking after a bad signal provides full insurance to protection buyers if and only if*

$$pAR \geq \pi\Delta\theta - (1-p)\,prob[\underline{s}]A\mathcal{P}. \tag{34}$$

In case of full insurance, the transfers are given by

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

As before in the case of risk-taking without margins and without clearing there is full insurance as in the first-best (as long as the amount of resources generated under risk-taking pAR is sufficiently high). Also as before, protection buyers must compensate protection sellers for the efficiency loss due to risk-taking so that the consumption of protection buyers falls short of the first-best level of consumption.

Condition (34) ensures that the limited liability constraints are slack under full insurance. On the left-hand side is the aggregate amount of 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. A sufficient condition for (34) to hold is

$$A \geq \pi \Delta \theta. \tag{35}$$

Risk-taking can be optimal only if it is relatively efficient, 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 (idiosyncratic) counterparty risk in the contract with risk-taking, while margins tackle ex-ante incentives in the contract with effort.

C. Margins and clearing: Effort or no effort?

The contract under which protection sellers exert effort after both signals leads to limited risk-sharing for buyers (Subsection III.A), while the contract with no effort after a bad signal (risk-taking) leads to full risk-sharing for protection buyers as long as protection sellers do not default (Subsection III.B). The next proposition characterizes the optimal choice between the two contracts as a function of the probability of success under risk-taking, p.

PROPOSITION 10: *Assume (35) 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 3. Consider the expected utility of a protection buyer when there is effort after a bad signal. It decreases when p increases. For this contract, 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 a protection buyer when there is no effort after a bad signal. In contrast with the previous case, it increases when p increases. For this contract, the only effect of an increase in p is to increase the amount of resources available after a bad signal. Hence the result, stated in the proposition, that risk-prevention effort after a bad signal is optimal if and only if p is lower than a threshold.

Insert Figure 3 here

As the probability of success under risk-taking, p, decreases, the pledgeability of a protection seller's assets increases. By Proposition 10, therefore, protection sellers with low pledgeable income enter contracts with risk-taking, while protection sellers with higher pledgeable income enter contracts with effort after both signals, with protection sellers with plegeable income above the threshold given in Lemma 1 offering ample risk-sharing without jeopardized incentives.

IV. Empirical implications

According to our theory, a strong and pledgeable asset base ($A\mathcal{P}$) helps maintaining protection sellers' risk-prevention incentives.[18] Asset pledgeability decreases with the cost of risk-prevention, the inefficiency of risk-management practices,[19] and the opacity and complexity of financial institutions and their activities. Our model predicts a non-linear, U-shaped relation between derivatives exposures and the pledgeability of assets (see the discussion following Proposition 10 and transfers in Propositions 1, 6, and 9).

Empirical implication 1. *Financial institutions with efficient risk-management and transparent activities optimally choose large derivatives exposures; financial institutions with less efficient risk-management and more opaque activities choose small derivatives exposures; financial institutions with very inefficient risk management and opaque activities choose large exposures (associated with significant counterparty risk).*

Derivatives exposure and protection sellers' incentives also depend 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-prevention effort (high p) can be viewed as a "benign"/low-risk economic situation. Derivatives contracts that offer ample insurance but undermine risk-management incentives will be traded in such a benign environment (see Proposition 10). This resonates with the notion that risk builds up in "good" times (see, e.g., Borio, 2011).

In this context, consider the effect of bad news. For example, when the underlying risk is that of mortgage defaults, declining house prices convey bad news. After bad news, protection sellers give up on risk-prevention. Hence, they become more likely to default. This creates correlation between mortgage values and the asset values of financial institutions without direct exposure to mortgage default.

An increase in the precision of the public information signal (λ) increases this endogenous correlation. Information about the performance of mortgage-backed securities and CDS contracts written on them was unavailable before 2006.[20] The ABX.HE indices providing this information were introduced only in January 2006. 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. Our theoretical analysis implies that the information then conveyed by the ABX.HE undermined the incentives of protection sellers. To the extent that ample insurance kept being written, it came at the expense of risk-taking. We summarize this discussion in our next empirical implication.

Empirical implication 2. *Derivatives contracts with large exposures are more likely to be underwritten when the economic environment seems benign. In this context, after bad news about the hedged risk, the expected value of the other assets of protection sellers decreases. The more accurate the information about the hedged risk, the stronger this contagion.*

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 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.

Empirical implication 3. *When risk-free rates are low compared to the return on productive investment opportunities, the optimal margin deposit is lower.*

In terms of alleviating the incentive problem, margins are particularly beneficial when the cost of risk-prevention effort is convex, and the optimal margin is higher the more convex risk-management costs are (see Section VI in the Online Appendix). Convexity in risk-management costs implies that the risk of each additional unit of assets is more costly to manage. This could be a feature of complex and opaque (information-sensitive) assets that require intense monitoring and information collection, which becomes more expensive as the size of assets under management increases. Convexity in risk-management cost could also be related to the liquidity of assets under management, with larger positions being more illiquid (e.g., due to a larger price impact and higher execution costs in case the position needs to be closed). The above discussion is summarized in our next implication:

Empirical implication 4. *The more risk-management costs increase with assets under management, the higher is the optimal margin.*

V. Policy implications

A. *Margins and equity capital*

We showed that margins allow for more incentive-compatible insurance as they ring-fence assets from protection seller moral hazard. Would capital requirements offer an alternative mechanism to reduce moral hazard? 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 margins and capital requirements. As argued below, our theoretical analysis implies that margins can be an attractive substitute to capital.

Margins reduce the need for equity capital. In our model, at $t = 0$, protection sellers have assets A and no liabilities. Hence, the book value of their equity capital (the difference between assets and liabilities) is A. Its market value, reflecting rationally anticipated future cash flows, is AR. 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

$$A[\alpha(\underline{s}) + (1 - \alpha(\underline{s}))\mathcal{P})], \tag{36}$$

to preserve protection seller's incentives to exert risk-prevention effort. Thus, the value of a protection seller's equity capital after a bad signal at $t = 1$ is

$$(1 - \alpha(\underline{s}))\,(R - \mathcal{P})\,A > 0, \tag{37}$$

which is the difference between $A\,[\alpha(\underline{s}) + (1 - \alpha(\underline{s}))R]$, the value of protection sellers' assets, and (36), the value of her liability. The interpretation is that the optimal contract with effort requires protection sellers to hold a minimum amount of equity (i.e., keep enough skin in the game) to make sure the incentive compatibility constraint holds.

Without margin calls (e.g., if there was no enforcement mechanism for margins), the incentive compatibility condition would be more demanding. Hence protection sellers would need to have a higher amount of equity (more skin in the game) to ensure that effort remains incentive compatible. In that sense, margins are a substitute to equity capital. 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. Consequently, margins allow protection sellers to engage in incentive compatible derivatives activity with less equity.

Higher capital is an alternative to margins, but can be infeasible. Another way to relax the incentive compatibility constraint after a bad signal would be to increase a protection seller's initial equity capital. This could be difficult to implement, however. In our simple agency-theoretic framework, raising capital from dispersed outside investors does not improve the incentives of the manager. On the contrary, it dilutes her ownership of the firm and reduces her incentives to exert effort. Thus, increasing capital relaxes incentive compatibility only if the additional capital belongs to the agent (increasing her skin in the game) or to investors closely monitoring the agent (reducing the severity of the moral hazard problem.) When these conditions cannot be met, margin requirements are more effective than capital requirements.

Moreover, margins, unlike equity, are linked to derivative positions. A margin call only occurs when the derivative position turns into a liability (which depends on information about the underlying asset). Capital requirements, independent of the development of derivative positions, could be wasteful as they would require equity capital even when derivative positions are profitable.

B. CCP design

The key advantage of the CCP over bilateral contracting is the mutualization of counterparty default risk. By insuring protection buyers, it makes them more eager to contract with protection sellers. At the same time, it makes each of them less eager to take costly actions to reduce protection sellers' default risk. Margin calls are one of the key instruments to reduce that risk. Thus, to implement the optimal contract characterized in this paper, one cannot delegate to the trading parties the task of designing their own individual margin calls. Such decentralization would lead to insufficient margining and excessive counterparty default. To see this, consider the case in which the optimal contract calls for effort even after bad news. Suppose the CCP offers the optimal transfers $\tau^S(\theta, \bar{s})$ and $\tau^B(\theta, \underline{s})$ described above, while letting each protection-seller/protection-buyer pair choose their own margin call. A limited-liability protection seller and a protection buyer insured against counterparty risk by the CCP both prefer to set $\alpha(s) = 0, \forall s$. But then the incentive compatibility condition for protection sellers does not hold and there will be excessive counterparty default risk. This is a form of free-riding, since the cost of that default is borne by all the other members of the CCP. To avoid such free-riding, margin calls must be mandated by the CCP.

VI. Conclusion

We analyze optimal contracts in the context of hedging with derivatives. We show how contracts designed to engineer risk-sharing can generate incentives for risk-taking. When the position of a 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 a protection seller. Thus, negative news about derivative positions can propagate to other lines of business of financial institutions and, when doing so, create endogenous counterparty risk.

When the sellers' 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. Such research in the context of derivatives and margins would complement Acharya and Viswanathan's (2011) analysis of fire-sale price externalities in the context of borrowers de-levering by selling assets.

References

Acharya, Viral V., and Alberto Bisin, 2014, "Counterparty Risk Externality: Centralized versus Over-the-counter Markets," *Journal of Economic Theory*, 149, 153–182.

Acharya, Viral V., Robert F. Engle, and Stephen Figlewski, Anthony W. Lynch and Marti G. Subrahmanyam, 2012, "Centralized Clearing for Credit Derivatives," in Viral V. Acharya, and Matthew Richardson ed.: *Restoring Financial Stability: How to Repair a Failed System* (New York University Stern School of Business).

Acharya, Viral V., and S. Viswanathan, 2011, "Leverage, Moral Hazard and Liquidity," *Journal of Finance*, 66 (1), 99–138.

Acharya, Viral V., and Tanju Yorulmazer, 2008, "Information Contagion and Bank Herding," *Journal of Money, Credit and Banking*, 40 (1), 215–231.

Allen, Franklin, and Elena Carletti, 2006, "Credit Risk Transfer and Contagion," *Journal of Monetary Economics*, 53 (1), 89–111.

Allen, Franklin, and Douglas Gale, 2000, "Financial Contagion," *Journal of Political Economy*, 108 (1), 1–33.

Berk, Jonathan B., and Richard C. Green, 2004, "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 112 (6), 1269–1295.

Biais, Bruno, Florian Heider, and Marie Hoerova, 2012, "Risk-sharing or Risk-taking? Counterparty Risk, Incentives and Margins," ECB Working Paper No. 1413.

Bolton, Patrick, and Martin Oehmke, 2011, "Credit Default Swaps and the Empty Creditor Problem," *Review of Financial Studies*, forthcoming.

Bolton, Patrick, and Martin Oehmke, 2013, "Should Derivatives be Senior?," *Journal of Finance*, forthcoming.

Bolton, Patrick, and David S. Scharfstein, 1990, "A Theory of Predation Based on Agency Problems in Financial Contracting," *American Economic Review*, 80 (1), 93–106.

Borio, Claudio E. V., 2011, "Rediscovering the Macroeconomic Roots of Financial Stability Policy: Journey, Challenges and a Way Forward," *Annual Review of Financial Economics,* 3, 87–117.

Brunnermeier, Markus K., and Lasse H. Pedersen, 2009, "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22 (6), 2201–2238.

Cifuentes, Rodrigo, Gianluigi Ferrucci, and Hyun S. Shin, 2005, "Liquidity Risk and Contagion," *Journal of the European Economic Association*, 3 (2–3), 556–566.

Dudley, William, 2014, "Global Financial Stability — the Road Ahead," remarks by the President and Chief Executive Officer of the Federal Reserve Bank of New York, at the Tenth Asia-Pacific High Level Meeting on Banking Supervision, 26 February 2014.

The Economist, 2012, "Centrally Cleared Derivatives: Clear and Present Danger," April 7.

Ellul, Andrew, and Vijay 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, Ingo, and Martin Scheicher, 2008, "The ABX: How Do the Markets Price Subprime Mortgage Risk?," *BIS Quarterly Review*, September, 67–81.

Fleming, Michael J., and Asani Sarkar, 2014, "The Failure Resolution of Lehman Brothers," *Economic Policy Review*, 20 (2).

Freixas, Xavier, Bruno M. Parigi, and Jean-Charles Rochet, 2000, "Systemic Risk, Interbank Relations, and Liquidity Provision by the Central Bank," *Journal of Money, Credit and Banking*, 32 (3), 611–638.

Froot, Kenneth A., Scharfstein, D., and J. Stein, 1993, "Risk Management: Coordinating Corporate Investment and Financing Policies," *Journal of Finance*, 48 (5), 1629–1658.

Gorton, Gary, and George 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.

Hirshleifer, Jack, 1971, "The Private and Social Value of Information and the Reward of Inventive Activity," *American Economic Review*, 61 (4), 561–574.

Holmström, Bengt, and Jean Tirole, 1997, "Financial Intermediation, Loanable Funds, and the Real Sector," *Quarterly Journal of Economics*, 112 (3), 663–691.

Holmström, Bengt, and Jean Tirole, 1998, "Private and Public Supply of Liquidity," *Journal of Political Economy*, 106 (1), 1–40.

Jensen, Michael C., and William H. Meckling, 1976, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3 (4), 305–360.

King, Mervyn A., and Sushil Wadhwani, 1990, "Transmission of Volatility between Stock Markets," *Review of Financial Studies*, 3 (1), 5–33.

Myers, Stewart C., 1977, "Determinants of Corporate Borrowing," *Journal of Financial Economics*, 5 (2), 147–175.

Parlour, Christine A., and Guillaume Plantin, 2008, "Loan Sales and Relationship Banking," *Journal of Finance*, 63 (3), 1291–1314.

Thompson, James R., 2010, "Counterparty Risk in Financial Contracts: Should the Insured Worry about the Insurer?," *Quarterly Journal of Economics*, 125 (3), 1195–1252.

Notes

[1]Total credit exposure is the sum total of net current credit exposure (NCCE) and potential future exposure (PFE). NCCE is the gross positive fair value of derivatives contracts less the dollar amount of netting benefits. PFE is an estimate of what the current credit exposure could be over time, based upon a supervisory formula in the risk-based capital rules.

[2]Thompson (2010) also assumes moral hazard of the seller of protection. But in his setting, moral hazard *facilitates* the provision of insurance.

[3]72% of the CDS AIG had sold by December 2007 were used by banks for capital relief (European Central Bank, 2009).

[4]See, e.g., Parlour and Plantin (2008) in the context of credit risk transfer in banking.

[5]The concavity of the objective function of the protection buyer can reflect institutional, financial or regulatory constraints, such as leverage constraints or risk-weighted capital requirements. For an explicit modeling of hedging motives see Froot, Scharfstein and Stein (1993).

[6]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.

[7]We show in the Online Appendix 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-prevention effort convex.

[8]Here effort improves returns in the sense of first-order stochastic dominance. In an earlier version of the paper (Biais, Heider and Hoerova, 2012) we showed that our results are robust when effort improves returns in the sense of second-order stochastic dominance, so that lack of effort corresponds to risk-shifting.

[9]It is indeed one the roles of market infrastructures to ensure such contractibility and enforceability.

[10]That assets with low information sensitivity are used as collateral is in line with Gorton and Pennacchi (1990).

[11]Margin calls before the signal s occurs are suboptimal because of the opportunity cost. Protection buyers are always better off waiting for the signal to occur. This way they keep the flexibility to call a margin after only one of the two realizations of the signal.

[12]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 does not alter our qualitative results.

[13]Without a contract protection sellers always exert effort since it is efficient to do so (see condition (1)).

[14]Note that $\lambda^*(p)$ is decreasing in p and $\lambda^* \to \frac{1}{2}$ as $p \to 1$. For reasonable values of p, the threshold λ^* is close to one half. For example, even for a relatively low p, $p = \frac{1}{2}$, we have $\lambda^* = 0.59$.

[15]Note however that instead of exogenous debt as in Myers (1977) our model involves endogenous liabilities pinned down by the optimal hedging contract.

[16]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.

[17]In our simple model this promised payment reflects a single trade. With multiple trades, the relevant expected payment would reflect the net exposure of protection sellers.

[18]While for simplicity protection sellers have no initial debt in our model, to gauge this implication empirically one should consider assets net of liabilities.

[19]Ellul and Yerramili (2013) propose a Risk Management Index measuring the organizational strength and independence of the risk management function within financial insititutions.

[20]Although the issuance of mortgage-backed securites was around \$ 2 trillion in every year from 2002 until 2006 (see, e.g., Fender and Scheicher, 2008).

Online Appendix: A. Proofs

Proof of Proposition 1 Form the Lagrangian using the objective (7), the feasibility constraints (6) with multiplier μ_{FC} and the participation constraint (8) with multiplier μ. For the moment we ignore the limited liability constraints (5) 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,

$$\text{prob}[\theta, s]u'(\theta + \tau^B(\theta, s)) - \mu_{FC}(\theta, s) = 0 \qquad \forall(\theta, s) \tag{A.1}$$

$$\mu\text{prob}[\theta, s] - \mu_{FC}(\theta, s) = 0 \qquad \forall(\theta, s). \tag{A.2}$$

Since marginal utility is strictly positive, it follows from (A.1) that $\mu_{FC}(\theta, s) > 0$ for all (θ, s) and hence the feasibility constraints bind. Since $\mu_{FC}(\theta, s) > 0$, it follows from (A.2) that the participation constraint binds. After substituting (A.1) into (A.2), 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

$$-(\text{prob}[\bar{\theta}, \bar{s}] + \text{prob}[\bar{\theta}, \underline{s}])\tau^B(\bar{\theta}, \bar{s}) - (\text{prob}[\underline{\theta}, \bar{s}] + \text{prob}[\underline{\theta}, \underline{s}])\tau^B(\underline{\theta}, \bar{s}) = 0 \tag{A.3}$$

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 $\tau^B(\underline{\theta}, \bar{s}) = \pi\Delta\theta$, 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) 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 (7), the limited liability constraints (5) with multipliers $\mu_{LL}(\theta, s)$, the participation constraint (8) with multiplier μ and the incentive compatibility constraints (10) with multipliers $\mu_{IC}(s)$.

In a bilateral contract, $\tau^S = -\tau^B$. Substituting for τ^S in the constraints, the first-order conditions of the Lagrangian with respect to $\tau^B(\theta, s)$ are given by:

$$u'(\theta + \tau^B(\theta, s)) = \mu + \frac{\mu_{LL}(\theta, s)}{\text{prob}[\theta, s]} + \frac{\mu_{IC}(s)}{\text{prob}[s]} \qquad \forall(\theta, s) \tag{A.4}$$

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

We conjecture that, given assumption (13), the limited liability constraints are slack in all states. We therefore proceed by solving the problem under the conjecture that $\mu_{LL}(\theta, s) = 0$ and then verifying (in the proof of Proposition 3) that this conjecture is correct.

We now show, by contradiction, that the participation constraint (8) binds. Suppose not. Plugging $\mu = 0$ and $\mu_{LL}(\bar{\theta}, s) = 0$ into (A.4) implies that $\mu_{IC}(s) > 0$ for all s. Hence, both incentive constraints bind, $-E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = s] = A\mathcal{P}$ for $s = \bar{s}, \underline{s}$. Therefore,

$$E[\tau^S(\tilde{\theta}, \tilde{s})] \equiv \text{prob}[\bar{s}]E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}] + \text{prob}[\underline{s}]E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] = -A\mathcal{P} < 0$$

But from the slack participation constraint, we have $E[\tau^S(\tilde{\theta}, \tilde{s})] > 0$, which is a contradiction. Hence, the participation constraint binds and $\mu > 0$.

Finally, we show that only the incentive constraint after a bad signal binds. Note that it cannot be that both incentive constraints are slack since we consider the case when 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 we used above to show that the participation constraint binds). We show by contradiction that it is the incentive constraint following a bad signal that binds. Suppose not and hence $\mu_{IC}\left(\underline{s}\right) = 0$. Equation (A.4) for $s = \bar{s}$ then implies that $u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s})) = u'(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}))$, and thus

$$\tau^B(\underline{\theta}, \bar{s}) - \tau^B(\bar{\theta}, \bar{s}) = \Delta\theta > 0. \tag{A.5}$$

Equation (A.4) for $s = \underline{s}$ implies $u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s})) = u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s}))$, and thus

$$\tau^B(\underline{\theta}, \underline{s}) - \tau^B(\bar{\theta}, \underline{s}) = \Delta\theta > 0 \tag{A.6}$$

Because $\mu_{IC}\left(\underline{s}\right) = 0$ and $\mu_{LL}(\underline{\theta}, \underline{s}) = \mu_{LL}(\underline{\theta}, \bar{s}) = 0$, it follows from (A.4) that $u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) \leq u'(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}))$ (as $\mu_{IC}\left(\bar{s}\right) \geq 0$), and thus

$$\tau^B(\underline{\theta}, \underline{s}) \geq \tau^B(\underline{\theta}, \bar{s}). \tag{A.7}$$

From the binding participation constraint

$$\text{prob}[\bar{s}]E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}] + \text{prob}[\underline{s}]E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = 0$$

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

$$E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}] > 0 \tag{A.8}$$

Using $\tau^B = -\tau^S$, (A.5) and (A.6)) we can write

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

Using $\tau^B = -\tau^S$ and (A.7) 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

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

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 (A.8). Consequently, the incentive constraint after a bad signal binds and the incentive constraint after a good signal must be slack. QED

Proof of Proposition 3 The optimal transfers are obtained by combining the binding participation constraint (8), the binding incentive constraint after a bad signal (10), $\tau^B = -\tau^S$, and full risk-sharing conditional on a signal (A.5) and (A.6).

We now check that, given our assumption (13), the limited liability constraints are slack in all states. We only need to check transfers when $\theta = \underline{\theta}$ since this is when a protection seller makes positive transfers, $-\tau^S = \tau^B > 0$. In state $(\underline{\theta},\underline{s})$, we have

$$\begin{aligned}\tau^B(\underline{\theta},\underline{s}) &= \underline{\pi}\Delta\theta + A\mathcal{P} < \underline{\pi}\Delta\theta + (\pi-\underline{\pi})\Delta\theta \qquad \text{[first-best not attainable]}\\ &= \pi\Delta\theta < AR. \qquad \text{[LL constraints slack in the first-best]}\end{aligned}$$

In state $(\underline{\theta},\bar{s})$, we need to show that

$$\tau^B(\underline{\theta},\bar{s}) = \bar{\pi}\Delta\theta - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A\mathcal{P} < AR.$$

Note that

$$\bar{\pi}\Delta\theta - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]}A\mathcal{P} < \bar{\pi}\Delta\theta - \frac{\text{prob}[\underline{s}]\,(1-p)\,A\mathcal{P}}{1-\text{prob}[\underline{s}]\,(1-\underline{\pi})\,(1-p)} < AR$$

where the last inequality is our assumption (13). The first inequality is equivalent to:

$$-\frac{1}{\text{prob}[\bar{s}]} < -\frac{(1-p)}{1-\text{prob}[\underline{s}]\,(1-\underline{\pi})\,(1-p)}$$

or, after simplifying,

$$\text{prob}[\underline{s}]\underline{\pi}\,(1-p) > -\,p$$

which is true since the left-hand side is positive while the right-hand side is negative. Hence, given (13), the limited liability constraint in state $(\underline{\theta}, \bar{s})$ is slack. QED

Proof of Proposition 4 Form the Lagrangian using the objective (16), the limited liability constraints (5) with multipliers $\mu_{LL}(\theta, s, R_j)$, the incentive compatibility constraint after a good signal ((10) with $\tilde{s} = \bar{s}$) and after a bad signal (17) with multipliers $\mu_{IC}(\bar{s})$ and $\mu_{IC}(\underline{s})$, respectively, and the participation constraint (18) with multiplier μ.

In a bilateral contract, $\tau^B(\theta, s, R_j) = -\tau^S(\theta, s, R_j)\ \forall(\theta, s, R_j)$, and the first-order conditions of the Lagrangian with respect to $\tau^B(\theta, s, R_j)$ are:

$$\begin{aligned}
u'(\theta + \tau^B(\theta, \bar{s}, R)) &= \mu + \frac{\mu_{LL}(\theta, \bar{s}, R)}{\text{prob}[\theta, \bar{s}]} + \frac{\mu_{IC}(\bar{s})}{\text{prob}[\bar{s}]} \qquad \theta = \bar{\theta}, \underline{\theta} && \text{(A.9)}\\
u'(\theta + \tau^B(\theta, \underline{s}, R)) &= \mu + \frac{\mu_{LL}(\theta, \underline{s}, R)}{p\text{prob}[\theta, \underline{s}]} - \frac{\mu_{IC}(\underline{s})}{p\text{prob}[\underline{s}]} \qquad \theta = \bar{\theta}, \underline{\theta} && \text{(A.10)}\\
u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, 0)) &= \mu + \frac{\mu_{LL}(\bar{\theta}, \underline{s}, 0)}{(1-p)\,\text{prob}[\bar{\theta}, \underline{s}]} + \frac{\mu_{IC}(\underline{s})}{(1-p)\,\text{prob}[\underline{s}]} && \text{(A.11)}
\end{aligned}$$

Note that $\tau^B(\underline{\theta}, \underline{s}, 0)$ is optimally set to zero.

We conjecture that, given assumption (13), the limited liability constraints are slack in all states, except in state $(\underline{\theta}, \underline{s})$ where $\tau^S(\underline{\theta}, \underline{s}, 0) = 0$ and the limited liability constraint binds. We therefore proceed by solving the problem under the conjecture that $\mu_{LL}(\theta, s, R_j) = 0$ and then verifying that this conjecture is correct.

First, note that the participation constraint binds. Suppose not, and hence $\mu = 0$. Since $\mu_{IC}(\underline{s}) \geq 0$, equations in (A.10) cannot hold. A contradiction.

Next, we show that the incentive constraint after a bad signal (17) is slack, implying $\mu_{IC}(\underline{s}) = 0$. Suppose that the constraint binds so that $A\mathcal{P} - \underline{\pi}\tau^S(\bar{\theta}, \underline{s}, 0) = -\underline{\pi}\tau^S(\bar{\theta}, \underline{s}, R) - (1-\underline{\pi})\tau^S(\underline{\theta}, \underline{s}, R)$ implying that

$$-\underline{\pi}\tau^S(\bar{\theta}, \underline{s}, R) - (1-\underline{\pi})\tau^S(\underline{\theta}, \underline{s}, R) \leq A\mathcal{P} \qquad \text{(A.12)}$$

since $-\tau^S(\bar{\theta}, \underline{s}, 0) \leq 0$.

Since the participation constraint binds, we have

$$\begin{aligned}\text{prob}[\underline{s}](1-p)A\mathcal{P}-\text{prob}[\underline{s}](1-p)\underline{\pi}\tau^S(\bar{\theta}, \underline{s}, 0) = \text{prob}[\bar{s}]\left[\bar{\pi}\tau^S(\bar{\theta}, \bar{s}, R) + (1-\bar{\pi})\tau^S(\underline{\theta}, \bar{s}, R)\right] \\ + \text{prob}[\underline{s}]p\left[\underline{\pi}\tau^S(\bar{\theta}, \underline{s}, R) + (1-\underline{\pi})\tau^S(\underline{\theta}, \underline{s}, R)\right]\end{aligned}$$

Using the binding incentive constraint (17) in the equation above and simplifying yields

$$\begin{aligned}&- \text{prob}[\bar{s}]\left[\bar{\pi}\tau^S(\bar{\theta}, \bar{s}, R) + (1-\bar{\pi})\tau^S(\underline{\theta}, \bar{s}, R)\right] \\ &- \text{prob}[\underline{s}]\left[\underline{\pi}\tau^S(\bar{\theta}, \underline{s}, R) + (1-\underline{\pi})\tau^S(\underline{\theta}, \underline{s}, R)\right] = 0\end{aligned} \tag{A.13}$$

Equations (A.12) and (A.13), respectively, imply that the optimal transfers $\tau(\bar{\theta}, \bar{s}, R)$, $\tau(\underline{\theta}, \bar{s}, R)$, $\tau(\bar{\theta}, \underline{s}, R)$ and $\tau(\underline{\theta}, \underline{s}, R)$ satisfy the incentive-compatibility condition inducing effort after bad news (10 when $\tilde{s} = \underline{s}$) and the participation constraint (8) in the contract with effort after both signals. Hence, inducing effort after both signals is feasible with these transfers. We now show that, given these transfers, the expected utility of the contract with effort after both signals is strictly higher than the expected utility of the contract without effort after bad news, i.e.:

$$\begin{aligned}&\pi\lambda u(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}, R)) + (1-\pi)(1-\lambda)u(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}, R)) + \pi(1-\lambda)u(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, R)) \\ &+ (1-\pi)\lambda u(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s}, R)) > \pi\lambda u(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}, R)) + (1-\pi)(1-\lambda)u(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}, R)) \\ &+ \pi(1-\lambda)[pu(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, R)) + (1-p)u(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, 0))] \\ &+ (1-\pi)\lambda[pu(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s}, R)) + (1-p)u(\underline{\theta})],\end{aligned}$$

or, equivalently, that

$$\begin{aligned}\pi(1-\lambda)\,(1-p)\left[u(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, R)) - u(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, 0))\right] \\ + (1-\pi)\lambda\,(1-p)\left[u(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s}, R)) - u(\underline{\theta})\right] > 0\end{aligned}$$

holds. It follows from equations (A.10) and (A.11) that $u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, R)) \leq u'(\bar{\theta} +$

$\tau^B(\bar{\theta}, \underline{s}, 0))$ and thus $\tau^B(\bar{\theta}, \underline{s}, R) \geq \tau^B(\bar{\theta}, \underline{s}, 0)$. Hence, the expression in the first square bracket above is non-negative. Next note that $\tau^B(\underline{\theta}, \underline{s}, R) > 0$. Suppose not. Equations (A.9) and (A.10) imply that $0 \geq \tau^B(\underline{\theta}, \underline{s}, R) \geq \tau^B(\underline{\theta}, \bar{s}, R) > \tau^B(\bar{\theta}, \bar{s}, R)$ and $0 \geq \tau^B(\underline{\theta}, \underline{s}, R) > \tau^B(\bar{\theta}, \underline{s}, R) \geq \tau^B(\bar{\theta}, \underline{s}, 0)$. But optimal transfers cannot be all zero or negative as the buyer would then get no insurance. Since $\tau^B(\underline{\theta}, \underline{s}, R) > 0$, the expression in the second square bracket above is positive. Hence, the protection buyer prefers to induce effort after bad news, contradicting the optimality of the contract with risk-taking after bad news. We conclude that if risk-taking after bad news is optimal, the incentive constraint after a bad signal (17) must be slack.

Under the assumption that the limited liability constraints are slack, (A.9)–(A.11) together with the slack incentive constraint after a bad signal imply that we have full sharing of the $\tilde{\theta}$ risk conditional on the signal except when the seller defaults in the $\underline{\theta}$ state: $u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}, R)) = u'(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s}, R))$ and $u'(\bar{\theta} + \tau(\bar{\theta}, \underline{s}, R)) = u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s}, R)) = u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}, 0))$. Therefore

$$\tau^B(\underline{\theta}, \tilde{s}) - \tau^B(\bar{\theta}, \tilde{s}) = \Delta\theta > 0 \tag{A.14}$$

We now show that the incentive constraint after a good signal ((10 when $\tilde{s} = \bar{s}$)) is also slack, implying $\mu_{IC}(\bar{s}) = 0$. When the constraint is slack, there is full insurance except when the seller defaults in $\underline{\theta}$ state, i.e. we have:

$$\tau^B(\tilde{\theta}, \bar{s}, R) = \tau^B(\tilde{\theta}, \underline{s}, R) \text{ and } \tau^B(\bar{\theta}, \underline{s}, R) = \tau^B(\bar{\theta}, \underline{s}, 0) \tag{A.15}$$

The optimal contract in this case is given by equations (A.14), (A.15) and the binding participation constraint. We now check under what conditions the incentive constraint following a good signal is indeed slack with that contract. Starting with the binding participation and using (A.14) and (A.15), we get

$$\begin{aligned} -\operatorname{prob}[\underline{s}](1-p)A\mathcal{P} = {} & \operatorname{prob}[\bar{s}][\tau^B(\underline{\theta}, \underline{s}, R) - \bar{\pi}\Delta\theta] \\ & + \operatorname{prob}[\underline{s}]p[\tau^B(\underline{\theta}, \underline{s}, R) - \underline{\pi}\Delta\theta] + (1-p)\operatorname{prob}[\underline{s}]\underline{\pi}\left[\tau^B(\underline{\theta}, \underline{s}, R) - \Delta\theta\right] \end{aligned}$$

Hence,

$$\tau^B(\underline{\theta}, \underline{s}, R) = \frac{\pi \Delta\theta - \text{prob}[\underline{s}]\,(1-p)\, A\mathcal{P}}{1 - \text{prob}[\underline{s}]\,(1-\underline{\pi})\,(1-p)} \tag{A.16}$$

For the incentive constraint following a good signal to be slack, it must be that

$$A\mathcal{P} > \bar{\pi}\tau^B(\bar{\theta}, \bar{s}, R) + (1-\bar{\pi})\tau^B(\underline{\theta}, \bar{s}, R) = \tau^B(\underline{\theta}, \underline{s}, R) - \bar{\pi}\Delta\theta$$

or, after substituting for $\tau^B(\underline{\theta}, \underline{s}, R)$ and simplifying,

$$A\mathcal{P} > \Delta\theta \frac{\pi - \bar{\pi}\left[1 - \text{prob}[\underline{s}]\,(1-\underline{\pi})\,(1-p)\right]}{1 + \text{prob}[\underline{s}]\underline{\pi}\,(1-p)} \tag{A.17}$$

Condition (A.17) is always satisfied if

$$\pi - \bar{\pi}\left[1 - \text{prob}[\underline{s}]\,(1-\underline{\pi})\,(1-p)\right] < 0 \tag{A.18}$$

since $A\mathcal{P} > 0$. Condition (A.18) is equivalent to $\lambda^2(1-p) - 2\lambda + 1 < 0$. This inequality holds under our assumption (12), i.e. for all $\lambda \geq \lambda^* \equiv \frac{1-\sqrt{p}}{1-p} > \frac{1}{2}$. This is because the left-hand side of the inequality above is decreasing in λ and it is equal to zero for λ^*.

We now check that, given our assumption (13), the limited liability constraints are slack in the optimal contract with no effort after a bad signal. First note that condition (A.18) is equivalent to

$$\frac{\pi}{\bar{\pi}} < 1 - \text{prob}[\underline{s}]\,(1-\underline{\pi})\,(1-p) \tag{A.19}$$

As the seller only makes positive transfers when $\theta = \underline{\theta}$ and $R_j = R$, we only need to check that the limited liability constraint for the transfer $\tau^B(\underline{\theta}, \tilde{s}, R)$ is slack, i.e.:

$$\frac{\pi \Delta\theta - \text{prob}[\underline{s}]\,(1-p)\, A\mathcal{P}}{1 - \text{prob}[\underline{s}]\,(1-\underline{\pi})\,(1-p)} < AR.$$

We have that

$$\begin{aligned}\frac{\pi\Delta\theta-\text{prob}[\underline{s}]\left(1-p\right)A\mathcal{P}}{1-\text{prob}[\underline{s}]\left(1-\underline{\pi}\right)\left(1-p\right)} &= \frac{\pi\Delta\theta}{1-\text{prob}[\underline{s}]\left(1-\underline{\pi}\right)\left(1-p\right)}-\frac{\text{prob}[\underline{s}]\left(1-p\right)A\mathcal{P}}{1-\text{prob}[\underline{s}]\left(1-\underline{\pi}\right)\left(1-p\right)}\\ &< \bar{\pi}\Delta\theta-\frac{\text{prob}[\underline{s}]\left(1-p\right)A\mathcal{P}}{1-\text{prob}[\underline{s}]\left(1-\underline{\pi}\right)\left(1-p\right)} \qquad \text{[using (A.19)]}\end{aligned}$$

By (13), we have that

$$\bar{\pi}\Delta\theta-\frac{\text{prob}[\underline{s}]\left(1-p\right)A\mathcal{P}}{1-\text{prob}[\underline{s}]\left(1-\underline{\pi}\right)\left(1-p\right)}<AR$$

so the limited liability constraint is slack. QED

Proof of Proposition 5 Form the Lagrangian using the objective (7), the feasibility constraints (6) with multiplier $\mu_{FC}\left(\theta,s\right)$, the limited liability constraints (5) with multipliers $\mu_{LL}\left(\theta,s\right)$, the feasibility constraints on margins (4) with $\mu_0\left(s\right)$ for $\alpha\left(s\right)\geq 0$ and $\mu_1\left(s\right)$ for $\alpha\left(s\right)\leq 1$, the participation constraint (19) with multiplier μ and the incentive compatibility constraints (20) with multipliers $\mu_{IC}(s)$.

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

$$\begin{aligned}\text{prob}[\theta,s]u'(\theta+\tau^B(\theta,s))-\mu_{FC}\left(\theta,s\right) &= 0 \qquad \forall(\theta,s) \quad \text{(A.20)}\\ \mu\text{prob}[\theta,s]+\mu_{LL}(\theta,s)+\text{prob}[\theta|s]\mu_{IC}(s)-\mu_{FC}\left(\theta,s\right) &= 0 \qquad \forall(\theta,s).\text{(A.21)}\end{aligned}$$

Since marginal utilities are positive, it follows from (A.20) that $\mu_{FC}\left(\theta,s\right)>0$ and hence all feasibility constraints bind:

$$\tau^B(\theta,s)=-\tau^S(\theta,s),\forall(\theta,s). \tag{A.22}$$

Using (A.20) to substitute for $\mu_{FC}\left(\theta,s\right)$ in (A.21) and rearranging, we obtain

$$u'(\theta+\tau^B(\theta,s))=\mu+\frac{\mu_{LL}(\theta,s)}{\text{prob}[\theta,s]}+\frac{\mu_{IC}(s)}{\text{prob}[s]} \qquad \forall(\theta,s) \tag{A.23}$$

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,

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

But since $R > \mathcal{P}$, this violates the incentive compatibility constraint (20) 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 (A.22), 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 (A.23) 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.

Next, we show by contradiction that the participation constraint (19) binds. Suppose not. Plugging $\mu = 0$ and $\mu_{LL}(\bar{\theta}, s) = 0$ (just shown above) into (A.23) 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,

$$E[\tau^S(\tilde{\theta}, \tilde{s})] = E[E[\tau^S(\tilde{\theta}, \tilde{s})|\tilde{s}]] = -E[\alpha(\tilde{s})A + (1 - \alpha(\tilde{s}))A\mathcal{P}] \tag{A.24}$$

From the participation constraint, we have

$$
\begin{aligned}
0 &\leq E[\tau^S(\tilde{\theta}, \tilde{s})] - E[\alpha(\tilde{s})]A(R - C - 1) \\
&= -E[\alpha(\tilde{s})A + (1 - \alpha(\tilde{s}))A\mathcal{P}] - E[\alpha(\tilde{s})]A(R - C - 1) \quad \text{[using (A.24)]} \\
&= -E[(1 - \alpha(\tilde{s}))A\mathcal{P} + \alpha(\tilde{s})A(R - C)].
\end{aligned}
$$

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 $\mu > 0$.

Finally, the first-order conditions of the Lagrangian from the proof of Proposition 5 with respect to $\alpha(s)$ are

$$
\frac{\mu_0(s) - \mu_1(s)}{A} + \mu_{IC}(s)(1 - \mathcal{P}) = \mu \text{prob}[s]\,(R - C - 1) + (R - 1)\mu_{LL}(\underline{\theta}, s) \qquad \forall s, \tag{A.25}
$$

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

The right-hand side of (A.25) is strictly positive since $R - C > 1$ and $\mu > 0$ (participation constraint binds). 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 6 We first show that the incentive constraint after a bad signal binds. Note that 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 5).

We now show by contradiction that the incentive constraint following a bad signal binds. Suppose not and hence $\mu_{IC}(\underline{s}) = 0$. In state $(\bar{\theta}, \bar{s})$, the limited liability constraint is slack (see proof of Proposition 5). In state $(\underline{\theta}, \bar{s})$, we conjecture that, given our assumption (13), the limited liability constraint is also slack. Therefore, $\mu_{LL}(\bar{\theta}, \bar{s}) = 0$ and $\mu_{LL}(\underline{\theta}, \bar{s}) = 0$. Equations (A.23) 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

$$\tau^B(\underline{\theta},\bar{s}) - \tau^B(\bar{\theta},\bar{s}) = \Delta\theta > 0. \tag{A.26}$$

After the bad signal, limited liability constraint in state $(\bar{\theta},\underline{s})$ is slack, $\mu_{LL}(\bar{\theta},\underline{s}) = 0$ (see proof of Proposition 5). 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 (A.23) 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

$$\tau^B(\underline{\theta},\underline{s}) - \tau^B(\bar{\theta},\underline{s}) = \Delta\theta > 0 \tag{A.27}$$

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 (A.23) that $u'(\underline{\theta} + \tau^B(\underline{\theta},\underline{s})) \leq u'(\underline{\theta} + \tau^B(\underline{\theta},\bar{s}))$, and thus

$$\tau^B(\underline{\theta},\underline{s}) \geq \tau^B(\underline{\theta},\bar{s}). \tag{A.28}$$

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

$$E[\tau^S(\tilde{\theta},\tilde{s})|\tilde{s} = \underline{s}] > 0. \tag{A.29}$$

Using full risk-sharing conditional on the signal (equations (A.26) and (A.27)) we can write

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

Using (A.28) and the binding feasibility conditions (A.22), 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

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

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 (A.29).

Now, consider the case when the limited liability constraint in state $(\underline{\theta}, \underline{s})$ binds. Since $\mu_{LL}(\bar{\theta}, \underline{s}) = 0$ and $\mu_{IC}(\underline{s}) = 0$, equations (A.23) 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

$$\tau^B(\underline{\theta}, \underline{s}) - \tau^B(\bar{\theta}, \underline{s}) \leq \Delta\theta. \tag{A.30}$$

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 (A.30) in conjunction with the binding feasibility constraints (A.22), we then have

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

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.

To obtain transfers after a good signal, combine the binding the participation constraint (19), the binding incentive constraint after a bad signal, and full risk-sharing after a good signal (A.26).

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 (A.27)). Using (A.27) and the binding incentive constraint after a bad signal, 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 the binding incentive constraint after a bad signal to obtain $\tau^B(\bar{\theta}, \underline{s})$.

Finally, we check that, under (13), the limited liability constraint in $(\underline{\theta}, \bar{s})$ is slack. Since $\alpha(\bar{s}) = 0$, the limited liability constraint (5) writes as $\tau^B(\underline{\theta}, \bar{s}) < AR$. Since $\tau^B(\underline{\theta}, \bar{s})$ decreases in $\alpha(\underline{s})$, we have that $\tau^B(\underline{\theta}, \bar{s}) < AR$ for all $\alpha(\underline{s})$ if and only if it is for $\alpha(\underline{s}) = 0$. This is the same transfer as in Proposition 3 and we showed that the limited liability constraint is slack in that case.

QED

Proof of Proposition 7 We first show 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 (A.23) for $s = \underline{s}$ imply that that there is full risk-sharing conditional on the bad signal. Hence, the individual transfers after the bad signal are given by (22) 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 (23). In particular, $\tau^B(\bar{\theta}, \underline{s}) = A > 0$. Equations (A.23) 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 (21), $\tau^B(\bar{\theta}, \bar{s}) < 0$. A contradiction.

We now derive the optimal margin after a bad signal, $\alpha^*(\underline{s})$. Using equations (A.23) to substitute for μ, $\mu_{IC}(\underline{s})$ and $\mu_{LL}(\underline{\theta}, \underline{s})$ in equation (A.25), we get

$$\begin{aligned}\frac{u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))} \quad = \quad & 1 + \frac{R - C - 1}{1 - \mathcal{P}} + \frac{\mu_1(\underline{s}) - \mu_0(\underline{s})}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))\text{prob}[\underline{s}]\,(1-\mathcal{P})\,A} \\ & + \frac{1-\underline{\pi}}{1-\mathcal{P}} \frac{u'(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})) - u'(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s}))}{u'(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s}))}\end{aligned} \tag{A.31}$$

where we used $\mu_{IC}(\bar{s}) = 0$ (since the incentive constraint is slack after a good signal).

Denote the right-hand side of (A.31) by φ. 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 $\frac{\partial \tau^B(\bar{\theta},\underline{s})}{\partial \alpha} = A\left[1 + \frac{(1-\underline{\pi})R - \mathcal{P}}{\underline{\pi}}\right] > 0$ since $R - \mathcal{P} > R - 1 > \underline{\pi}(R-1)$). Hence, φ is decreasing in α. If $\varphi(0) < 1 + \frac{R-C-1}{1-\mathcal{P}}$, then

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

for any $\alpha \in [0,1]$ (since the last term is non-negative). By equation (A.31) 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 (A.27)), and the last term in equation (A.31) 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

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

We now check under what condition the limited liability constraint in state $(\underline{\theta},\underline{s})$ is slack. Using (22), we have that the constraint is slack if and only if:

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

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 is slack if and only if

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

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 (22)) implying that

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

In case the limited liability constraint binds, it also must be that $\tau^B(\bar{\theta},\underline{s})<0$. This is because equations (23) imply that

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

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

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

or, equivalently,

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

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

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

QED

Proof of Proposition 8 Form the Lagrangian using the objective (29), the feasibility constraints (30) and (31) with multipliers $\mu_{FC}(\theta,s)$, the limited liability constraints (5) with multipliers $\mu_{LL}(\theta,s,R)$, the feasibility constraints on margins

(4) with $\mu_0(s)$ for $\alpha(s) \geq 0$ and $\mu_1(s)$ for $\alpha(s) \leq 1$, the incentive compatibility constraints (20) with multipliers $\mu_{IC}(s)$ and the participation constraint (33) with multiplier μ.

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

$$\text{prob}[\theta, s]u'(\theta + \tau^B(\theta, s)) - \mu_{FC}(\theta, s) = 0 \qquad \forall(\theta, s) \tag{A.32}$$

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

$$\begin{aligned}
\mu\text{prob}[\theta, \bar{s}] + \mu_{LL}(\theta, \bar{s}, R) + \text{prob}[\theta|\bar{s}]\mu_{IC}(\bar{s}) - \mu_{FC}(\theta, \bar{s}) &= 0 \quad \forall(\theta, \bar{s}, R) &\text{(A.33)}\\
\mu\text{prob}[\theta, \underline{s}] + \frac{\mu_{LL}(\theta, \underline{s}, R)}{p} - \text{prob}[\theta|\underline{s}]\frac{\mu_{IC}(\underline{s})}{p} - \mu_{FC}(\theta, \underline{s}) &= 0 \quad \forall(\theta, \underline{s}, R) &\text{(A.34)}\\
\mu\text{prob}[\theta, \underline{s}] + \frac{\mu_{LL}(\theta, \underline{s}, 0)}{1-p} + \text{prob}[\theta|\underline{s}]\frac{\mu_{IC}(\underline{s})}{1-p} - \mu_{FC}(\theta, \underline{s}) &= 0 \quad \forall(\theta, \underline{s}, 0) &\text{(A.35)}
\end{aligned}$$

Since marginal utilities are positive, it follows from (A.32) that $\mu_{FC}(\theta, s) > 0$ and hence the feasibility constraints (30) and (31) bind.

Using (A.32) to substitute for $\mu_{FC}(\theta, s)$ in (A.33)-(A.35) and rearranging, we obtain

$$\begin{aligned}
u'(\theta + \tau^B(\theta, \bar{s})) &= \mu + \frac{\mu_{LL}(\theta, \bar{s}, R)}{\text{prob}[\theta, \bar{s}]} + \frac{\mu_{IC}(\bar{s})}{\text{prob}[\bar{s}]} \quad \forall(\theta, \bar{s}, R) &\text{(A.36)}\\
u'(\theta + \tau^B(\theta, \underline{s})) &= \mu + \frac{\mu_{LL}(\theta, \underline{s}, R)}{p\text{prob}[\theta, \underline{s}]} - \frac{\mu_{IC}(\underline{s})}{p\text{prob}[\underline{s}]} \quad \forall(\theta, \underline{s}, R) &\text{(A.37)}\\
u'(\theta + \tau^B(\theta, \underline{s})) &= \mu + \frac{\mu_{LL}(\theta, \underline{s}, 0)}{(1-p)\,\text{prob}[\theta, \underline{s}]} + \frac{\mu_{IC}(\underline{s})}{(1-p)\,\text{prob}[\underline{s}]} \quad \forall(\theta, \underline{s}, 0) &\text{(A.38)}
\end{aligned}$$

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

Combining (A.37) and (A.38) yields

$$(1-p)\,\mu_{LL}(\theta, \underline{s}, R) - p\mu_{LL}(\theta, \underline{s}, 0) = \text{prob}[\theta|\underline{s}]\mu_{IC}(\underline{s}) \qquad \forall(\theta, \underline{s}) \tag{A.39}$$

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 (20) 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 (30), 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 (A.36) 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 (33) binds. Suppose not, i.e. $\mu = 0$. Using $\mu = 0$ in (A.37), 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 (A.36), 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,

$$\alpha(\underline{s})A + (1 - \alpha(\underline{s}))A(R - \mathcal{P}) = -\underline{\pi}\tau^S(\bar{\theta}, \underline{s}, 0) - (1 - \underline{\pi})\tau^S(\underline{\theta}, \underline{s}, 0) \tag{A.40}$$

If the limited liability constraints (5) for $R_j = 0$ 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 (A.40) is strictly smaller than $\alpha(\underline{s})A$. Since $(1 - \alpha(\underline{s}))A(R - \mathcal{P}) \geq 0$, the left-hand side of (A.40) is greater or equal to $\alpha(\underline{s})A$. A contradiction. If the limited liability constraints (5) for $R_j = 0$ 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 (33), 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

$$\text{prob}[\bar{s}][\alpha(\bar{s})A(R - C) + (1 - \alpha(\bar{s}))A\mathcal{P}] + \text{prob}[\underline{s}]A[(1 - p)\mathcal{P} + pR] \leq 0 \tag{A.41}$$

Since both terms on the right-hand side of (A.41) 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 (A.39) 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

$$-\tau^S(\bar{\theta}, \underline{s}, R) = -\tau^S(\underline{\theta}, \underline{s}, R) = \alpha(\underline{s})A + (1 - \alpha(\underline{s}))AR \tag{A.42}$$

Using (A.42) in the incentive constraint after a bad signal (32) 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 (A.39) 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 (33), we get

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

Simplifying yields

$$\text{prob}[\bar{s}]\alpha(\bar{s})A(R - C - 1) + \text{prob}[\underline{s}]ApR + \text{prob}[\underline{s}](1 - p)A\mathcal{P} = \text{prob}[\bar{s}]E[\tau^S(\theta, \bar{s}, R)] \tag{A.43}$$

For equation (A.43) to hold, it must be that $E[\tau^S(\theta, \bar{s}, R)] > 0$. By the binding feasibility constraint (30), 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 (A.43). 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 (A.36) and $\mu_{LL}(\bar{\theta}, \underline{s}, R) > 0$ and $\mu_{IC}(\underline{s}) = 0$ in (A.37), 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:

$$\begin{aligned}
\tau^B(\bar{\theta},\bar{s}) &> \tau^B(\bar{\theta},\underline{s}) = -p\tau^S(\bar{\theta},\underline{s},R) - (1-p)\tau^S(\bar{\theta},\underline{s},0) \text{ [binding feasibility constraint]} \\
&= p\left[\alpha(\underline{s})A + (1-\alpha(\underline{s}))AR\right] + (1-p)\alpha(\underline{s})A \quad \text{[binding LL constraints in } \underline{s}\text{]} \\
&= \alpha(\underline{s})A + p(1-\alpha(\underline{s}))AR > 0 \qquad \text{(A.44)}
\end{aligned}$$

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 (A.44), this implies that $E[\tau^B(\theta,\bar{s},R)] > 0$, a contradiction with (A.43). 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 (A.44), this implies that $E[\tau^B(\theta,\bar{s},R)] = -E[\tau^S(\theta,\bar{s},R)] > 0$, a contradiction with (A.43).

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 μ_{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 (A.39), 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 (A.39), it follows that $\mu_{LL}(\bar{\theta},\underline{s},0) > 0$. Hence,

$$\tau^B(\bar{\theta},\underline{s}) = p\left(\alpha(\underline{s})A + (1-\alpha(\underline{s}))AR\right) + (1-p)\alpha(\underline{s})A \qquad \text{(A.45)}$$

Using $\mu_{LL}(\bar{\theta},\underline{s},R) > 0$ and $\mu_{LL}(\underline{\theta},\underline{s},R) = 0$ in (A.37), 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

$$\tau^B(\bar{\theta},\underline{s}) < \tau^B(\underline{\theta},\underline{s}) \qquad \text{(A.46)}$$

must hold. However, we also have that

$$\begin{aligned}
\tau^B(\underline{\theta},\underline{s}) &= -p\tau^S(\underline{\theta},\underline{s},R) - (1-p)\tau^S(\underline{\theta},\underline{s},0) \qquad \text{[binding feasibility constraint]} \\
&\leq p\left(\alpha(\underline{s})A + (1-\alpha(\underline{s}))AR\right) + (1-p)\alpha(\underline{s})A \qquad \text{[LL constraints]} \\
&= \tau^B(\bar{\theta},\underline{s}) \qquad \text{[using (A.45)]}
\end{aligned}$$

which contradicts (A.46). 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 (A.39).

QED

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

$$\frac{\mu_0(\bar{s}) - \mu_1(\bar{s})}{A} + \mu_{IC}(\bar{s})(1-\mathcal{P}) = \mu \text{prob}[\bar{s}](R - C - 1) + (R-1)\mu_{LL}(\underline{\theta}, \bar{s}, R) \tag{A.47}$$

$$\mu_{LL}(\underline{\theta}, \underline{s}, 0) + \frac{\mu_0(\underline{s}) - \mu_1(\underline{s})}{A} = \mu \text{prob}[\underline{s}](pR - 1) + (R-1)\mu_{LL}(\underline{\theta}, \underline{s}, R) \tag{A.48}$$

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 (A.47) 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 (A.39) 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 (A.48) yields

$$\frac{\mu_0(\underline{s}) - \mu_1(\underline{s})}{A} = (pR - 1)\left[\mu \text{prob}[\underline{s}] + \frac{\mu_{LL}(\underline{\theta}, \underline{s}, R)}{p}\right] \tag{A.49}$$

If $pR \geq 1$, then the right-hand side of (A.49) 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 (A.49) 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 7), implying that the optimal contract with effort is strictly preferred to the contract with risk-taking and $\alpha\left(\underline{s}\right) = 1$.

QED

Proof of Proposition 9 The optimal transfers follow from asserting full risk-sharing across all states and using the binding participation constraint. Condition (34) 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 (20) we have:

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

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 8).

QED

Proof of Proposition 10 We first show that for $p < \max\left\{\frac{R-C-1}{R-1}, \frac{1}{R}\right\}$ 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 (35) 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 8, the contract with effort strictly dominates the contract with risk-taking in this case.

We now consider the case when $p \geq \max\left\{\frac{R-C-1}{R-1}, \frac{1}{R}\right\}$. 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 6, the expected utility of the protection buyer under effort is given by

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

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

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

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 7). Due to the binding incentive constraint after a bad signal (Proposition 6), 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 6, the expected utility of the protection buyer is given by

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

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

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

where we used $\frac{\pi(1-\lambda)}{\underline{\pi}}$ =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 7), $1-\alpha^*(\underline{s})>0$. Using (A.23), the fact that the limited liability constraints in states $(\bar{\theta},\bar{s})$ and $(\bar{\theta},\underline{s})$ are always slack and that the incentive constraint after a bad signal binds (Proposition 6), we have that $u'\big(\bar{\theta}+\tau(\bar{\theta},\underline{s})\big) > u'\big(\bar{\theta}+\tau(\bar{\theta},\bar{s})\big)$ 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 9 in (29), 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

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

Lastly, note that as $p \to \frac{R-C}{R}$ (or, equivalently, as $\mathcal{P} \to 0$), 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 $p < \max\left\{\frac{R-C-1}{R-1}, \frac{1}{R}\right\}$, the contract with effort is optimal. For $p \to \frac{R-C}{R}$, the contract with risk-taking is optimal. For $\max\left\{\frac{R-C-1}{R-1}, \frac{1}{R}\right\} \leq p < \frac{R-C}{R}$, 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

Online Appendix: B. Extensions and robustness

Renegotiation

The optimal contract that induces effort after both good and bad news is contingent on the signal s. One may wonder whether the optimal outcome could also be achieved by renegotiating at time 1, after s is observed, an initial contract that is independent of the signal.

For brevity and simplicity, rather than offering a general treatment of this question, we discuss the underlying economic forces in the context of an example. Suppose we start from an initial contractual transfer $\tau^B(\theta)$ independent of the signal. For example, suppose we take it to be the transfer prevailing in the optimal contingent contract after good news $\tau^B(\theta, \bar{s})$. Would both parties (protection buyer and protection seller) agree to switch from $\tau^B(\theta) = \tau^B(\theta, \bar{s})$ to $\tau^B(\theta, \underline{s})$ after observing bad news?

First consider protection sellers. Sticking to $\tau^B(\theta, \bar{s})$ after a bad signal violates her incentive compatibility constraint. Thus, she does not exert risk-prevention effort and fails with probability $1 - p$. Her expected gain then is:[21]

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

If instead she switches to $\tau^B(\theta, \underline{s})$, and thus exerts risk-prevention effort, she expects to obtain:

$$\underline{\pi} \left(AR - \tau^B(\bar{\theta}, \underline{s}) \right) + (1 - \underline{\pi}) \left(AR - \tau^B(\underline{\theta}, \underline{s}) \right) - AC. \tag{B.2}$$

By switching, the protection seller increases the expected payoff on her assets. She also reduces the payment to the protection buyers as $\tau^B(\underline{\theta}, \underline{s}) < \tau^B(\underline{\theta}, \bar{s})$. Thus, switching is quite attractive for her, as we now establish more formally. Substituting for the transfers and re-arranging, (B.2) is larger than (B.1) if and only if

$$A\mathcal{P} < E[\theta] - \text{prob}[\bar{s}] E[\tilde{\theta} | \underline{s}] \tag{B.3}$$

which is satisfied whenever first-best is not reachable (see (11)).

Now turn to protection buyers. Sticking to $\tau^B(\theta, \bar{s})$ after a bad signal implies higher transfers, but undermines the incentives of protection sellers. When the CCP insures against counterparty risk, a protection buyer does not internalize the cost of default of his counterparty. Consequently, a protection buyer does not accept to switch from the initial contract to $\tau^B(\theta, \underline{s})$ after bad news. Thus, the simple renegotiation game we proposed does not implement the optimal contract. This negative result extends to a larger class of renegotiation games. To the extent that they are insured against counterparty risk, investors are not willing to downscale initially generous insurance promises to preserve incentives. This suggests that, with centralized clearing, the adjustment of transfers, contingent on the arrival of information, should be factored in the initial contract.

What if, instead, trading occurs bilaterally and there is no centralized clearing? Then, in the simple renegotiation game proposed above, after observing bad news the protection buyer knows he will be exposed to counterparty risk if he sticks to $\tau^B(\theta, \bar{s})$. In this case his expected utility is

$$\underline{\pi} u(\bar{\theta} + \tau^B(\bar{\theta}, \bar{s})) + (1 - \underline{\pi}) p u(\underline{\theta} + \tau^B(\underline{\theta}, \bar{s})). \tag{B.4}$$

If instead he switches to $\tau^B(\theta, \underline{s})$, the protection buyer's expected utility is

$$\underline{\pi} u(\bar{\theta} + \tau^B(\bar{\theta}, \underline{s})) + (1 - \underline{\pi}) u(\underline{\theta} + \tau^B(\underline{\theta}, \underline{s})). \tag{B.5}$$

Substituting for the transfers, (B.5) is larger than (B.4) if and only if

$$p < \frac{\frac{u(E(\theta|\underline{s}) + A\mathcal{P})}{u\left(E(\theta|\bar{s}) - \frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]} A\mathcal{P}\right)} - \underline{\pi}}{1 - \underline{\pi}}. \tag{B.6}$$

If (B.6) holds, i.e., if effort strongly improves productive efficiency, then although they started from an initial non-contingent contract $\tau^B(\theta)$, both parties are happy to switch to $\tau^B(\theta, \underline{s})$ after bad news, in order to preserve incentives. This suggests that, without centralized clearing, initially non-contingent contracts could, in some cases, be successfully renegotiated to the optimal contract.

On the other hand, if p is relatively large and (B.6) does not hold, protection buyers don't find it very attractive to renegotiate to lower insurance payments after bad news. In that case, renegotiation is unlikely to implement the optimal contract.

Derivative payoffs

The payoff from an interest rate swap is symmetric, while 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. Does this skewness in the payoff have an effect on incentives?

To analyze the effect of an increase in the skewness of the hedged risk on incentives formally, we increase the probability π of a good outcome for protection buyers' risk θ while keeping its mean and the standard deviation constant.[22] An increase of π 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 counterveiling effect when the skewness π 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. Thus, at high levels of π, a further increase of skewness mutes the negative effect of bad news on incentives.[23] But, as long as $\pi < \lambda$ (the precision of the signal s), the negative effect on incentives from larger amounts of risk dominates and more skewness leads to more severe incentive problems. In this case, it is more difficult to maintain risk-management incentives when the underlying risk is skewed.

Non-linear cost of risk-prevention effort

Up to now, we assumed that the cost of risk-prevention effort increases linearly in the assets under management. We now relax this assumption and allow the cost of effort to be convex in assets under management.[24] This reflects the notion that, while controlling and preventing risk is relatively easy when the amount of assets under management is low, it gets more complex and costly when this amount is large.[25] Thus, we assume that the cost of risk-prevention effort is equal to

$$c(1-\alpha)A + \gamma(1-\alpha)^2 A^2 \tag{B.7}$$

when assets under management are $(1-\alpha)A$.

In the analysis above we had $\gamma = 0$. If the cost is convex, $\gamma > 0$, there is a new effect: as margins increase, assets under management decrease, and so does the marginal cost of risk-prevention. We hereafter analyze the optimal contract arising in this case. Since margins do not play any role for centrally cleared contracts with risk-taking, we only need to consider the contract with effort. As in Section III.A, the feasibility and participation constraints bind: there is no reason to have idle resources or to leave rents to protection sellers. 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. As in Section III.A, the incentive compatibility condition after bad news simplifies to

$$\alpha(\underline{s})A + (1-\alpha(\underline{s}))A\mathcal{P}(\alpha(\underline{s})) \geq E[\tau^B(\tilde{\theta},\tilde{s})|\tilde{s} = \underline{s}]. \tag{B.8}$$

The difference with Section III.A is that now the pledgeable return depends on the size of the margin call after a bad signal:

$$\mathcal{P}(\alpha(\underline{s})) \equiv R - \frac{c + \gamma(1-\alpha(\underline{s}))A}{1-p}. \tag{B.9}$$

Margins improve risk-sharing when they relax the incentive constraint after a bad signal, i.e., when the left-hand-side of (B.8) is increasing in α, i.e., when

$$\mathcal{P}(\alpha(\underline{s})) - (1-\alpha(\underline{s}))\,\mathcal{P}' < 1. \tag{B.10}$$

It is easier to satisfy condition (B.10) when the cost of effort is convex (so that $\mathcal{P}' > 0$) than when it is linear (and $\mathcal{P}' = 0$). This reflects the above mentioned effect that as margins increase, the marginal cost of effort decreases.

To determine the optimal margin with convex costs, we proceed as in Section III.A. The transfers τ^B and τ^S have the same structure as in Proposition 6, except that $\mathcal{P}$ is now given by (B.9). We obtain the following proposition (where, as before, φ denotes the ratio of the marginal utility of a protection buyer after a bad and a good signal).

PROPOSITION B.1: *With a convex cost of risk-prevention effort, $\gamma > 0$, an optimal interior margin after a bad signal $\alpha^*(\underline{s})$ is given by*

$$\varphi(\alpha^*(\underline{s})) = 1 + \frac{R - C - 1}{1 - [\mathcal{P}(\alpha^*(\underline{s})) - (1 - \alpha^*(\underline{s}))\,\mathcal{P}']}. \tag{B.11}$$

Proof of Proposition B.1: 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 (7) as

$$U = \text{prob}[\bar{s}]u(E[\tilde{\theta} + \tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]) + \text{prob}[\underline{s}]u(E[\tilde{\theta} + \tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]). \tag{B.12}$$

Using the binding incentive and participation constraints, 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 $\alpha(\underline{s})$):

$$\frac{\partial U}{\partial \alpha} = \text{prob}[\bar{s}]\frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \bar{s}]}{\partial \alpha}\bar{u}' + \text{prob}[\underline{s}]\frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \alpha}\underline{u}', \tag{B.13}$$

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

$$\frac{\partial E[\tau^B(\tilde{\theta}, \tilde{s})|\tilde{s} = \underline{s}]}{\partial \alpha} = A\left[1 - \mathcal{P}(\alpha) + (1 - \alpha)\mathcal{P}'(\alpha)\right]. \tag{B.14}$$

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

$$X \equiv 1 - \mathcal{P}(\alpha) + (1-\alpha)\,\mathcal{P}'(\alpha) \tag{B.15}$$

The derivative is positive if and only if $X > 0$. This is condition (B.10) in the text.

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

$$\frac{\partial E[\tau^B(\tilde{\theta},\tilde{s})|\tilde{s}=\bar{s}]}{\partial \alpha} = -\frac{\text{prob}[\underline{s}]}{\text{prob}[\bar{s}]} A\left[(R-C-1)+X\right] \tag{B.16}$$

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 (B.14), (B.15) and (B.16), we can write (B.13) 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). When $\gamma \geq 0$, then $1 > R - \frac{c}{1-p}$ is sufficient for $X > 0$ for all α.

QED

As in Proposition 7, the optimal interior margin reflects the trade-off between improved risk-sharing across signals and the opportunity cost of margin deposits. But now $\mathcal{P}' > 0$, which lowers the right-hand-side of the inequality. Holding $\mathcal{P}$ fixed (to reason other things equal), this increases the value of $\alpha^*(\underline{s})$ (the solution of (B.11)). Thus, we obtain the following comparative static result.

PROPOSITION B.2: *Other things equal, the greater is the convexity of the cost of risk-prevention, the larger is the optimal margin.*

Proof of Proposition B.2: 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 α^* 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 (B.14), (B.15) and (B.16), the cross-partial derivative becomes

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

Moreover,

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

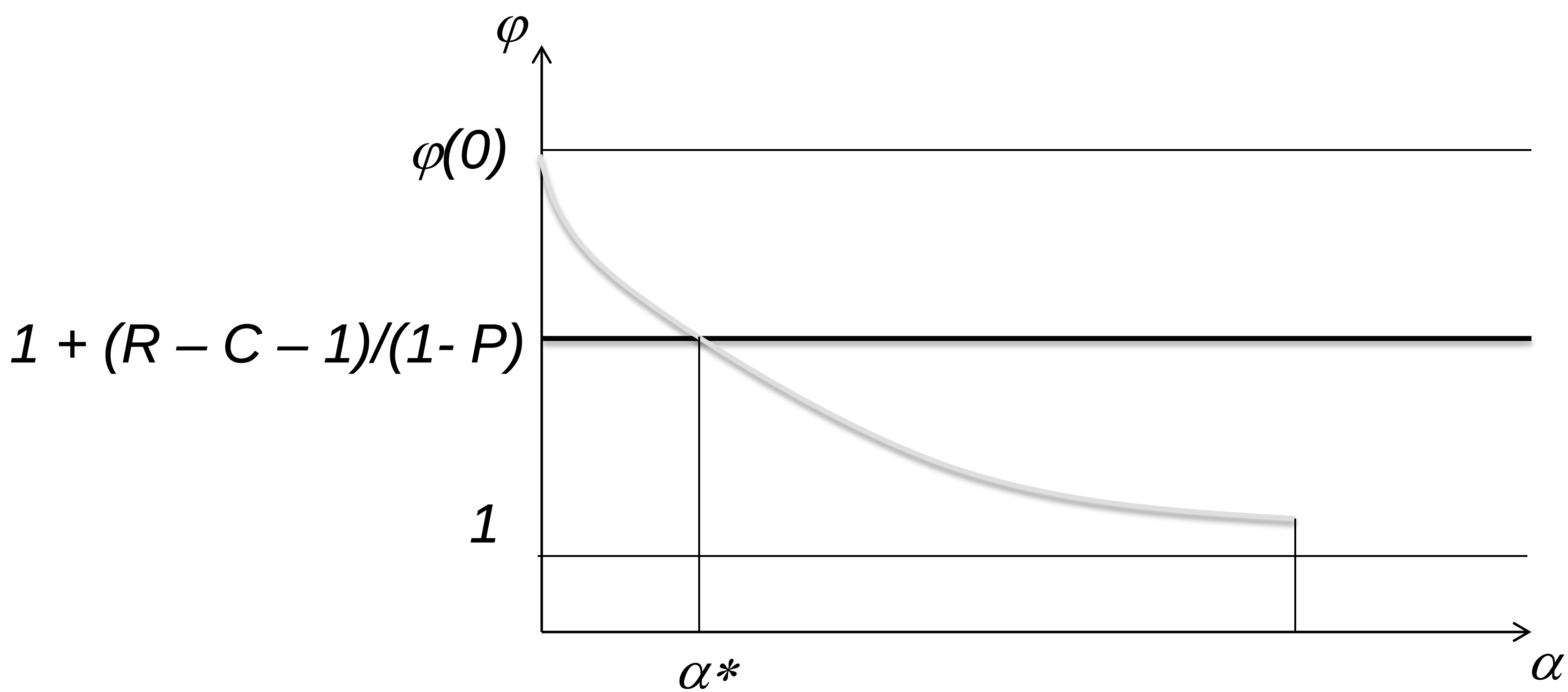
When $\gamma \geq 0$ then α^* is a local maximum and $R - \frac{c}{1-p} < 1$ is sufficient for $X > 0$. And when $X > 0$, the cross-partial derivative is positive.

QED

Figure 1: Timing

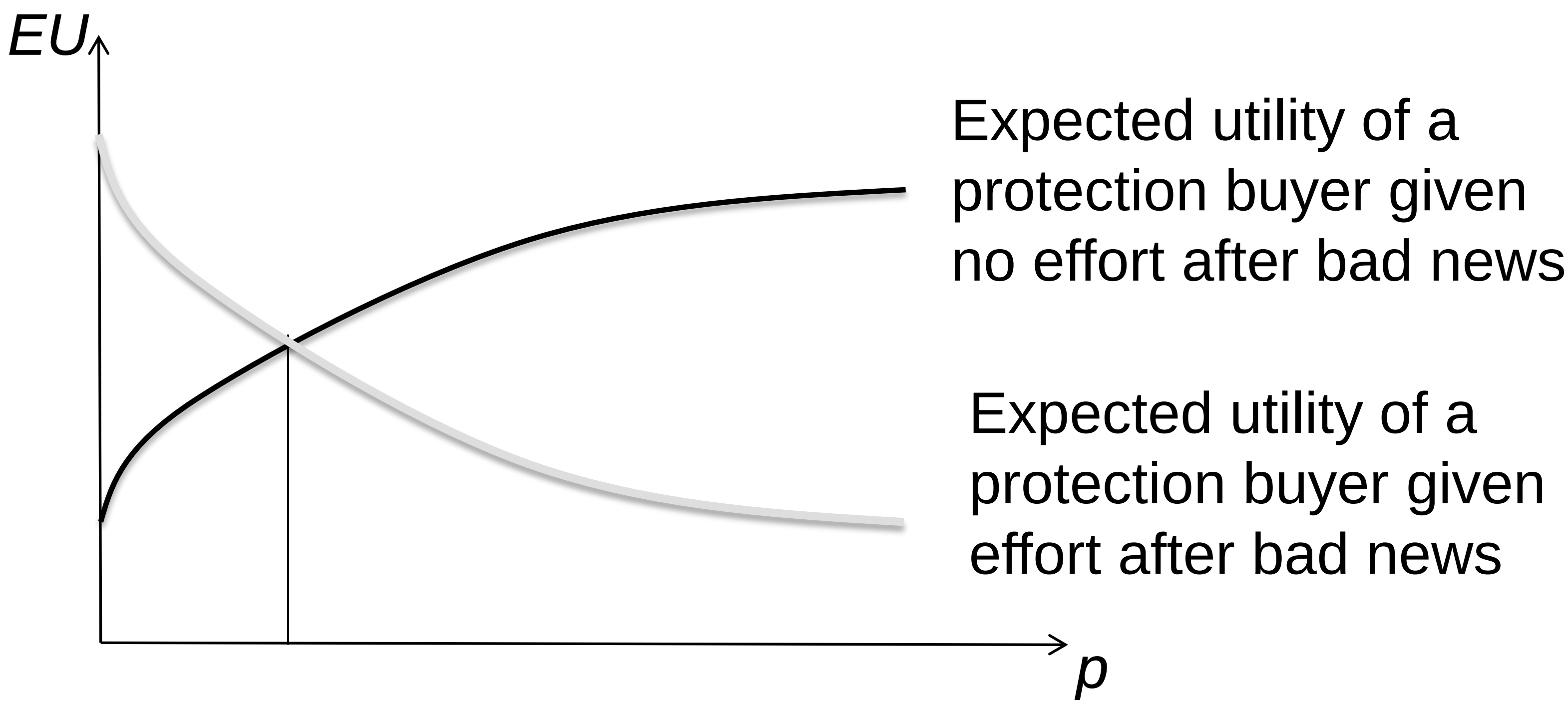
$t = 0$	$t = 1$	$t = 2$
Optimal contract design	Signal s about buyer's risk θ observed by all	Buyer's risk θ & seller's risk R_j realize
Participation decisions	Margin call $\alpha(s)$	Transfers τ^S τ^B (contingent on all observables: s, θ, R_j)
	Effort or not	

Figure 2: Optimal margins



Notes: Choice of the optimal margin $\alpha*$. The ratio (φ, y-axis) of the marginal utility of a protection buyer after bad and good news (grey curve) as a function of the margin after bad news (α, x-axis). The trade-off between the cost and benefit of a margin is given by *1 + (R – C – 1)/(1- P)* (black line). The numerator of the fraction, R-C-1, is the opportunity cost of depositing a margin; the denominator, 1-P, is the incentive benefit of a margin.

Figure 3: Optimal effort level



Notes: Choice of the optimal effort level. Expected utility (EU, y-axis) of a protection buyer under effort after bad news (grey curve) and under no effort after bad news (black curve) as a function of the probability of success under risk-taking (p, x-axis). Probability denotes a cutoff level of p beyond which the contract with no effort after bad news is optimal.