Optimal Incentives and Securitization of Defaultable Assets *

Semyon Malamud[†], Huaxia Rui[‡], and Andrew Whinston[§]

(forthcoming at *Journal of Financial Economics*)

Abstract

We study optimal securitization in the presence of an initial moral hazard. A financial intermediary creates and then sells to outside investors defaultable assets, whose default risk is determined by the unobservable costly effort exerted by the intermediary. We calculate the optimal contract for any given effort level and show the natural emergence of extreme punishment for defaults, under which investors stop paying the intermediary after the first default. With securitization contracts optimally designed, we find securitization improves the intermediary's screening incentives. Furthermore, the equilibrium effort level and the surplus converge to their first best levels with sufficiently many assets.

Keywords: securitization, mortgage-backed securities, asset-backed securities, moral hazard, default risk

JEL Classifications: D8, G21, G24, G32

---

*The research of S. Malamud was supported in part by NCCR FINRISK, Project A5. We thank an anonymous referee for very useful comments and suggestions that allowed us to significantly improve the paper. We also thank Darrell Duffie, Stefano Giglio, Tomasz Piskorski, René Stulz, Alexei Tchistyi, David Thesmar, Nancy Wallace, and participants at the CEPR Asset Pricing Meeting and UC Berkeley Real Estate Seminar for useful discussions. Existing errors are our sole responsibility.

[†]Swiss Finance Institute, EPF Lausanne. e-mail: semyon.malamud@epfl.ch

[‡]University of Texas at Austin, McCombs School of Business. e-mail: huaxia@utexas.edu

[§]University of Texas at Austin, McCombs School of Business and Department of Economics, e-mail: abw@uts.cc.utexas.edu

1. Introduction

The disastrous meltdown of the structured securitization market during the 2007-2008 financial crisis has been largely attributed to insufficient regulation and misalignment of incentives between market participants. Many policy-makers have argued that, under the "originate-and-distribute" business model, intermediaries did not bear the credit risk of borrower default, which led to deterioration in the credit quality of the underlying assets. Several policy proposals have been made to ensure that securitizers have strong incentives to monitor the quality of the assets they package. Some of these proposals were ultimately included in the Dodd-Frank Wall Street Reform and Consumer Protection Act, including a requirement that securitizers retain at least five percent of the default risk of assets in an asset-backed security (ABS).

Although this five percent retention rule is supposed to ensure that the intermediaries have "skin in the game," whether this rule is efficient is not at all clear.[1] Furthermore, the degree of its inefficiency might depend in a very non-trivial way on the precise nature of the default risk of the securitized assets. The theoretical foundations of the five percent rule have their origins in the literature on *static* optimal security design, where it has been shown that simple retention rules efficiently resolve the problem of informational asymmetry (see, e.g., DeMarzo and Duffie, 1999; Diamond and Rajan, 2000). However, this simple rule may fail to take into account of the intrinsic *dynamic long-term* nature of ABS default risk. The fact that contractual payments can be conditioned only on infrequent, discrete default events might completely alter the structure of the optimal incentive provision. These effects have been essentially ignored in the literature on optimal security design until a recent paper by Hartman-Glaser, Piskorski, and Tchistyi

[1]This potential inefficiency has been a topic of extensive discussions by policy makers. Several improvements have been suggested, such as permitting securitizers to select a form of risk retention from a menu of options. See, e.g., Luis A. Aguilar, "Speech by SEC Commissioner: Realigning Incentives in the Securitization Market," U.S. Securities and Exchange Commission, March 30, 2011, http://www.sec.gov/news/speech/2011/spch033011laa-item-1.htm

(2011) (henceforth, HPT (2011)). They were the first to introduce a general framework for studying optimal securitization of defaultable assets and formalize the major economic mechanisms that are important for incentive provision. In this paper, we build on the insights of HPT (2011) and study the following general questions: (1) How can we design efficient incentive alignment mechanisms taking into account of the dynamic nature of ABS default risk? (2) How does this optimal design depend on the market conditions and the nature of the assets underlying the ABS? (3) To what extent are natural market mechanisms able to induce efficient incentive alignment, and how can regulation improve these mechanisms? These are the questions we address in this paper.

As in HPT (2011), we consider the problem of dynamic optimal contracting between an intermediary (the securitizer) and outside investors in the presence of an initial moral hazard. Given a form of the contractual agreement, the intermediary optimally chooses the level of costly effort that he exerts initially to screen the assets that will be securitized.[2]

We use the standard Black and Cox (1976) structural model to model default risk of any single asset in the securitized basket. In this model, the borrower's distance to default is characterized by a stochastic process that fluctuates over time, and the default occurs when this process falls below a given threshold. This process can be interpreted as operational cash flows or firm value (when the borrower is a firm), income (when the borrower is an individual), or house price (in the case of a mortgage). The default risk of a borrower is then characterized by three directly interpretable economic parameters: initial distance to default (determined by how much the stochastic process has to fall for the default to occur), growth rate, and volatility. These three degrees of freedom allow us to investigate the effects of these three sources of default risk on the structure

[2]The cost of exerting the effort may be interpreted in two ways: First, it is the direct cost of carefully monitoring the potential borrowers; second, it is the indirect cost of missing the opportunity to securitize low-quality (high-risk) assets. This indirect cost may be significantly higher than the simple direct screening cost.

of the optimal contract between the intermediary and investors. We characterize the optimal contract in closed form and show that, under certain circumstances, it can exhibit surprising patterns such as large payments to the intermediary after a fixed number of defaults occur one immediately after another (*paying for default cascades*).[3] However, when the desired effort level implements the lowest default hazard rate and the risk aversion of the market participants is sufficiently small, we show that the optimal contract exhibits *extreme punishment for defaults:* It makes positive transfers to the intermediary only until the time of the first default. In particular, this result shows that the optimal contract of HPT (2011) is robust and holds even in the presence of risk aversion, provided the latter is sufficiently small. Furthermore, we also show that a lumpy contract is optimal in the risk neutral limit, in agreement with the findings of HPT (2011).

To assess the effect of bargaining power on the optimal contract design, we compare two polar cases: the competitive case, in which the intermediary has all the bargaining power in designing the contract, and the monopolistic case, in which the investor has all the bargaining power in designing the contract. Naturally, we might expect that the ability to sell all the credit risk "off the balance sheet" reduces the securitizer's screening incentives. On the other hand, when investors have all the bargaining power, we might expect that it would be easier for them to provide correct incentives for the intermediary. This leads us to formulate the following two conjectures:

Conjecture 1. Securitization leads to lax screening if the securitizer has all the bargaining power.

Conjecture 2. The optimal screening effort when the securitizer has all the bargaining power is lower than the optimal screening effort when the investor has all the bargaining power.

[3]As we discuss in the main text, this contract structure has the undesirable feature of being quite vulnerable to manipulation by the intermediary.

Surprisingly, our analytical results indicate that the following statements are true:

- *Conjecture 1 never holds* when risk aversions of both agents are sufficiently small. In this case, *securitization always improves the intermediary's screening incentives,* independent of who has the bargaining power.

- *Conjecture 2 generally does not hold.* In fact, we show that, when risk aversions are small, the equilibrium effort level of the intermediary does not depend on the bargaining power allocation. Reducing the severity of the moral hazard problem through pooling is always in the best interest of the intermediary; pooling allows the intermediary to achieve surplus extraction, which is arbitrarily close to the first best. Furthermore, for some parameter values, the equilibrium effort level will be higher when the intermediary designs the contract than when the investor designs the contract.

The economic intuition behind these surprising findings comes from the fact that, in our model: (i) The effort is priced (because investors have rational expectations and correctly anticipate the intermediary's optimal effort level for any given contract); and (ii) the cost of additional effort (relative to the no securitization case) is always less than the discount imposed due to the fact that the effort is priced. Part (i) is clear: Indeed, even if the intermediary designs a contract that leaves it with a minimal exposure to the underlying default risk, investors know that, with such contract form, the intermediary will optimally choose a low effort. Therefore, investors will demand a high premium for the increased default risk. Part (ii) is not at all obvious and is driven by the risk neutrality of the market participants and the relative impatience of the intermediary. Indeed, if all market participants are risk neutral and the investor discounts future cash flows at a higher rate than the intermediary, the incremental benefit of effort is always

greater for the investor than for the intermediary.[4]

This premium effectively reduces the intermediary's cost of effort and might make exerting a higher effort optimal. When investors have all the bargaining power, providing incentives for the intermediary to reduce the overall default risk certainly is in their interest. However, the effective cost that investors have to pay for this incentive provision might be higher in some circumstances than the effective effort cost reduction when the securitizer has all the bargaining power. In this case, the equilibrium effort level in the competitive case is higher than that in the case of a monopolistic investor.

When agents are risk neutral, we show that, surprisingly, the optimal contracts in the two polar cases differ only in the payment they make at time zero. In particular, the equilibrium effort level is independent of the bargaining power allocation and converges to the first best level when the number of securitized assets is sufficiently large. Increasing the minimal number of assets in the securitized pool always improves incentives and facilitates surplus extraction.

Our results have potential implications for securitization practices and regulation. In contrast to the conventional wisdom, we argue that securitization may improve incentives and therefore increase overall welfare of the society. This can happen even in situations when intermediaries are completely unregulated and have full bargaining power. We believe that properly designed securitization contracts may significantly reduce the "lemons spread"[5] and therefore improve the efficiency of these ABS markets. Finally, we note that, in our model, we make an implicit assumption that no information about asset

[4]The presence of risk aversion may potentially change the effect of securitization. For example, consider a model in which the intermediary is risk averse and the investor is risk neutral, and effort only reduces riskiness of cash flows, but does not change their present value. Since effort is costly and the investor is risk neutral, any efficient contract will implement the lowest effort. However, in the absence of securitization, it will always be optimal for the intermediary to exert effort if the cost of effort is not too high. We thank an anonymous referee for suggesting with this very illustrative example.

[5]See, e.g. Downing, Jaffee and Wallace (2009) who find that this spread accounts for up to 45% of the overall prepayment spread of mortgage-backed securities.

quality is verifiable. Under this assumption, credit rating agencies are essentially redundant and the degree of the credit risk of an ABS is determined solely by the structure of the corresponding contract and can therefore be directly inferred by rational investors. By contrast, if a credit rating agency could provide credible information about the asset quality, default contingent contracts may become unnecessary. The relative efficiency of the two different mechanisms depends on how costly it is for the rating agencies to verify the information, relative to how costly it is to provide incentives through default contingent contracts. Nevertheless, introducing optimal contracts into standard securitization practices would reduce the role of ratings in securitization, which is particularly important given the obvious failure of rating agencies to provide credible information during (and before) the crisis (Bolton, Freixas, and Shapiro, 2011).

We discuss the related literature in the following paragraphs.

Several empirical papers study the effect of misaligned incentives on the ABS market during the 2007-2008 crisis. Mian and Sufi (2009) find evidence that the extraordinary subprime mortgage growth from 2002 to 2005 was driven by a sharp rise in securitization, and they further suggest that the moral hazard associated with securitization might have caused the high mortgage default rates during that period and contributed to the global financial crisis. Keys, Mukherjee, Seru, and Vig (2010) exploit a specific rule of thumb in the mortgage lending market to examine whether the securitization process reduces the incentives of financial intermediaries to screen borrowers carefully, and their results suggest that existing securitization practices adversely affected the screening incentives of subprime lenders. Using comprehensive sales data of mortgage-backed securities from 1991 through 2002, Downing, Jaffee, and Wallace (2009) also show that securitized assets are of low quality with unfavorable performance. Thus, all three papers find empirical support for Conjecture 1 above. The discrepancy between our theoretical predictions and the empirical evidences can be explained by two factors: (1) Investors in the ABS

market were not sophisticated enough (i.e., did not have rational expectations) to fully understand the intermediaries' incentives issue; and (2) the contracts used in the industry were far from being optimal or efficient in any sense. As we discuss above, resolving these two issues might significantly improve the functionality of the securitization market.

Ashcraft and Schuermann (2009), Fender and Mitchell (2009), and Kane (2009) provide a detailed discussion of the chain of incentive conflicts that led to the subprime mortgage crisis. Minton, Stulz and Williamson (2009) and Stulz (2010) study how the adverse effect of the use of credit derivatives on lender's incentives contributed to the financial crisis.[6] In particular, Stulz (2010, p.90) argues that: "Rather than blaming derivatives markets, such as the credit default swap market, for being too large, it might make as much sense to regret that derivatives markets were not larger." We believe that our theoretical results strongly support this argument. Larger, more efficient, and better designed markets for securitizing and sharing credit risk might significantly improve social welfare.[7]

Our paper is also clearly related to theoretical literature on optimal security design. One large strand of this literature studies *static* optimal security design in the presence of asymmetric information. DeMarzo and Duffie (1999) develop a model in which the issuer has private information about the future payoff and signals a high-value security by its willingness to retain a portion of the issue. They study the problem of ex ante security design: The issuer designs the security before obtaining a signal about its value. DeMarzo and Duffie show that, under certain conditions, the optimal ex ante security design is a standard debt. DeMarzo (2005) studies whether pooling and tranching is optimal for an informed security issuer. Biais and Mariotti (2004) extend the Duffie

[6]See also Partnoy and Skeel (2006) for a pre-crisis warning.

[7]Of course, strict regulation of the credit risk exposure of market participants participating in true sale transactions is necessary for efficient incentive alignment. For example, according to the Dodd-Frank Act, a securitizer is prohibited from evading the risk retention requirements by hedging or transferring the credit risk that they are required to retain.

and DeMarzo (1999) model and study how securities and issuance mechanisms can be designed to mitigate the adverse effect of market imperfections on liquidity. They study separately the competitive case and the monopolistic case, just as we do in our model.

Gorton and Pennachi (1990) and Boot and Thakor (1993) show that, when both informed and uninformed investors are present in the market, it is optimal to split the asset into two securities: one senior and less information-sensitive security, and one junior and more information-sensitive security. Fulghieri and Lukin (2001) and Axelson (2007) study the security design problem that arises when outside investors have private information about the firm. Each of these papers, however, assumes that the unknown quality of the underlying assets is exogenous and is not affected by the effort of the seller; therefore, risk retention in these models is a signaling device. In contrast, investors in our model rationally anticipate the quality of the assets for any given contract. Thus, asymmetric information is present ex-ante because of moral hazard, but it is absent ex-post. This distinction is also true in the one-period optimal contracting problem with moral hazard, studied by Innes (1990). Another large strand of the literature is motivated by spanning risks. For surveys, see Allen and Gale (1994) and Duffie and Rahi (1995).

Substantial literature addresses dynamic optimal contracting with a repeated moral hazard, in which an agent makes a choice about effort in every period, and the current effort choice affects only the current outcome. See DeMarzo and Sannikov (2006), DeMarzo and Fishman (2007), and Sannikov (2008) for this stream of literature. In contrast, the moral hazard problem in our paper is *persistent:* A single action of the securitizer at time zero determines the probability distribution of all future cash flows. Furthermore, the nature of information flow in our model is very unique because new information arrives only at default events.

The closest to our work is the recent paper by Hartman-Glaser, Piskorski, and

Tchistyi (2011), which is the first paper to derive an optimal securitization contract in a dynamic setting. Their analysis generates new insights regarding the dynamic nature of the contracting problem and their optimal contract takes a an elegant and simple form: It makes a single payment to the intermediary after a waiting period if default occurs during this period. The derivation of this result in HPT (2011) is based on the following assumptions: (1) both the intermediary and the investor are risk neutral; (2) default times are exponentially distributed; and (3) it is optimal for the investors to provide incentives for the intermediary to exert the highest effort level. We show that the result of HPT (2011) is robust and holds in more general settings, up to small modifications. Namely, assuming that the risk aversions of both agents are sufficiently small and that higher effort leads to a lower default hazard rate, we show that the optimal contract is characterized by *multiple waiting periods,* where the intermediary gets partially remunerated for the absence of defaults during every subsequent waiting period. This result is very general and holds for a large class of default time distributions and multiple effort levels. As an application, we use the benchmark Black and Cox (1976) model to study how different sources of default risk interact and how they affect the form of the optimal contract.

2. Model Setup

We consider a continuous time optimal contracting problem between two agents: an intermediary S (the seller) and an outside investor B (the buyer). At time $t = 0$, the intermediary creates a pool of N defaultable assets (e.g., issues loans or mortgages, or acquires defaultable bonds) and sells this pool to the investor. The quality of the assets in the pool depends on the intermediary's unobservable (and hence non-contractible) costly effort e that can take a finite number of values, $e \in \{e_1, \cdots, e_K\}$ with $e_1 < \cdots < e_K$. The direct utility cost of exerting effort level e_j is equal to C_j , $j = 1, \cdots, K$. In the

case of a binary effort ($K = 2$), we write $e \in \{e_L, e_H\}$, where e_L corresponds to low effort and e_H to high effort.

As in HPT (2011), we assume that, conditional on the effort level e_j, exerted by the intermediary at time 0, the random default times $\{T_1, \cdots, T_N\}$ of the N assets in the pool are independent and identically distributed[8] with the cumulative distribution and the survival functions given by:

$$F_{e_j}(t) \equiv \text{Prob}[T_i < t \,|\, e_j] \qquad \text{and} \qquad G_{e_j}(t) \equiv \text{Prob}[T_i \geq t \,|\, e_j] = 1 - F_{e_j}(t) \tag{1}$$

for all $i \in \{1, \cdots, K\}$. Consistent with intuition, we say that an effort reduces default risk if it reduces the probability of default of any given asset i over any given time horizon $[0, t]$. Everywhere in the sequel, we make the following assumption:

Assumption. We assume that higher effort reduces default risk—that is $F_{e_j}(t)$ is monotone decreasing in $j \in \{1, \cdots, K\}$ for any $t \geq 0$. Furthermore, we assume that $F_{e_j}(t)$ has a strictly positive, continuous density $p_{e_j}(t)$ for any $j \in \{1, \cdots, K\}$.

For simplicity, we assume that each securitized asset $i = 1, \cdots, N$ is a defaultable bond (e.g., loan, mortgage, etc.) paying a fixed coupon rate u until the the default occurs at time T_i, and a recovery coupon rate $R < u$ after the default.[9] Let D_t denote the total number of defaults that have occurred up to time t. Then, because individual default times are i.i.d., the information set of the investor and the intermediary coincides with

[8] We do not analyze correlated defaults and the effects of systemic risk on the shape of the optimal contract. These effects may be important both for securitization of consumer loans and mortgages, where default risk is directly related to home prices, and for securitization of defaultable bonds, where defaults may be frailty-correlated. See Duffie, Eckner, Horel, and Saita (2009). Our methods can be directly extended to this more general setting, and we leave it as a topic for future research.

[9] It would be more realistic to assume that the bond holder gets a lump sum payment $\rho\, u/r$, equal to the value of the bond, multiplied by the recovery rate ρ. By picking $R = \rho u$, we get that the present value of the bond after default is equal to $\rho u/r$, and the two assumptions thus are equivalent if the agents can invest the recovery payment into a risk-free bond. We make the continuous payment assumption for technical reasons because the agents are maximizing utility from continuous consumption.

the filtration $\mathcal{F}_t$ generated by the process D_t. Let

$$\tau_n \;=\; \inf\{t \,>\, 0 :\; D_t \geq n\}$$

be the stopping time of the n-th default, and we set $\tau_0 = 0$. Then, the total payment rate d_t of the pool is given by the stochastic process:

$$d_t \;=\; (N - D_t)\, u \;+\; D_t\, R \;=\; \sum_{n=0}^{N} \delta_n\, \mathbf{1}_{t\in[\tau_n,\tau_{n+1})}$$

where

$$\delta_n \;=\; (N-n)u \,+\, n\, R\,. \tag{2}$$

A securitization contract specifies a transfer schedule from the investor to the intermediary, contingent on the history of defaults. Namely, the schedule is given by a sequence of payment rates $\{x_n(t, \tau_{[1,n]}),\ n \geq 0\}$ specifying the rate $x_n \geq 0$ *(limited liability for the intermediary)* that the investor transfers to the intermediary at the instant of time t after exactly n defaults have occurred, conditional on their occurrence times $\tau_1, \cdots, \tau_n$. Both the intermediary (agent S) and the investor (agent B) are risk averse,[10] and they maximize the discounted intertemporal expected utilities u_S and u_B from their life-time consumption, discounted at the rates γ and r respectively. As is common in the literature on dynamic optimal contracting, we assume that the intermediary is relatively impatient–that is, $\gamma \;>\; r$.[11]

For each $i = S, B$ the utility function u_i is assumed to be strictly increasing and concave; it is defined on an interval $(\ell_i, +\infty)$ for some $\ell_i \in [-\infty, 0)$ and satisfies the

[10]Later on, we consider the risk neutral limit case when the risk aversions of both agents converge to zero.

[11]See, e.g., DeMarzo and Duffie (1999), DeMarzo and Sannikov (2006), and HPT (2011). This assumption is typically justified as a preference for cash or for additional investment opportunities by the agent (intermediary), as in DeMarzo and Duffie (1999).

standard Inada conditions

$$\lim_{c\downarrow \ell_i} u_i'(c) = +\infty \ , \ \lim_{c\to+\infty} u_i'(c) = 0\,.$$

The utility of the intermediary from entering the contract $\{x_n\}$ after exerting an effort e_j is given by[12]

$$U_S(\{x_n\}, e_j) \ \equiv \ E\left[\int_0^\infty e^{-\gamma\, t}\, u_S(x_{D_t}(t, \tau_{[1,D_t]}))\, dt \,|\, e_j\right] \ - \ C_j \tag{3}$$

and the corresponding utility of the investor is given by:

$$U_B(\{x_n\}, e_j) \ \equiv \ E\left[\int_0^\infty e^{-r\, t}\, u_B(d_t \ - \ x_{D_t}(t, \tau_{[1,D_t]}))\, dt \,|\, e_j\right]\,. \tag{4}$$

We can now formulate the optimal contracting problem and describe efficient allocations, corresponding to the two polar cases: the competitive case, in which the intermediary has all the bargaining power in designing the contract, and the monopolistic case, in which the investor has all the bargaining power in designing the contract.

Following the lines of HPT (2011)[13], we make the following definition.

Definition 2..1 *An efficient allocation is a quadruple* $(\tilde{U}_B^0, \tilde{U}_S^0, \{x_n\}, e_j)$ *consisting of*

- *a contract* $\{x_n\}$ *with* $x_n \geq 0$ *for all* $n \geq 0$;
- *an effort level* e_j;
- *the utility* $\tilde{U}_S^0 \equiv U_S(\{x_n\}, e_j)$ *for the seller after entering the contract;*

[12]The assumption that the agent (intermediary) derives utility from the continuous flow of consumption, offered by the principal (investor) is standard in the optimal contracting literature. See, e.g., Sannikov (2008).

[13]See, HPT (2011), Definition 1. We thank an anonymous referee for suggesting this form of definition.

- *the utility* $\tilde{U}_B^0 \equiv U_B(\{\delta_n - x_n\}, e_j)$ *of the buyer after entering the contract,*

where the pair $(\{x_n\}, e_j)$ *fulfills the optimality condition*

$$\tilde{U}_B^0 \;=\; \max_{\{(y_n, e_k)\}} U_B(\{\delta_n - y_n\}, e_k) \tag{I}$$

where the maximization is over all contract-effort pairs $(\{y_n\}, e_k)$ *satisfying the limited liability (LL), incentive compatibility (IC) and the intermediary's participation constraints (PC):*

$$y_n \;\geq\; 0 \qquad \textit{for all} \;\; n \geq 0, \tag{LL}$$

$$e_j \;=\; \arg\max_{e_k} U_S(\{y_n\}, e_k), \tag{IC}$$

$$U_S(\{y_n\}, e_k) \;\geq\; \tilde{U}_S^0\,. \tag{PC}$$

This definition describes the entire Pareto frontier of contracting outcomes. Denote by $\mathcal{E}$ the set of efficient allocations. In this paper we will study two particular efficient allocations, corresponding to extreme bargaining power allocations. Namely, the monopolistic investor case and the competitive investor case.

- The equilibrium in the monopolistic case corresponds to the efficient allocation for which[14]

$$\tilde{U}_B^0 \;=\; \max\{U :\; (U, \tilde{U}_S^0, \{x_n\}, e_j) \in \mathcal{E} \text{ for some } \{x_n\}, e_j \text{ and } \tilde{U}_S^0 \geq U_S^0\}$$

- the equilibrium in the competitive case corresponds to the efficient allocation with

[14]Interestingly enough, as HPT (2011) show, (PC) does not bind in the monopolistic case if the outside option U_S^0 is sufficiently small.

$\tilde{U}_B^0 = U_B^0$ that maximizes $\tilde{U}_S^0$. That is,[15]

$$\tilde{U}_S^0 \;=\; \max\{U:\; (U_B^0, U, \{x_n\}, e_j) \in \mathcal{E} \text{ for some } \{x_n\}, e_j\}\,.$$

As is common in the literature, we will solve the optimal contracting problem in two steps: First, find the the optimal contract, implementing any given effort level; second, find the optimal (equilibrium) effort level.

3. The Optimal Contract for a Given Effort Level

In this section, we fix an effort level e_j, characterize the optimal contract implementing this effort level, and study its properties.

By definition, the stopping times $\tau_1 < \cdots < \tau_N$ coincide with the order statistics of the individual default times $T_1, \cdots, T_N$. In particular, $\tau_1 = \min\{T_k, k = 1, \cdots, N\}$ and $\tau_N = \max\{T_k, k = 1, \cdots, N\}$. We will denote by $f_k^{e_j}(\tau_{[1,k]})$ the joint density of

$$\tau_{[1,k]} \;\equiv\; (\tau_1, \cdots, \tau_k),\; k \leq N$$

conditional on the effort level e_j.[16]

Fix an effort level e_j. Denote by $\text{Prob}^{e_j}[\tau_{k+1} > t|\tau_k]$ the probability that the $(k+1)$-th default occurs not earlier than at time $t \geq \tau_k$.[17] Since the intermediary receives $x_k(t, \tau_{[1,k]})$ only if $(k+1)$-th default occurs not earlier than at time $t \geq \tau_k$, for the time period between the k-th and the $(k+1)$-th default events, the gain for the intermediary

[15] It may happen that there are multiple efficient allocations corresponding to $\tilde{U}_B^0 = U_B^0$.

[16] Lemma A.1 in the Appendix provides an explicit expression for this density.

[17] It is possible to show that $\text{Prob}^{e_j}[\tau_{k+1} > t|\tau_k] \;=\; \frac{(G_{e_j}(t))^{N-k}}{(G_{e_j}(\tau_k))^{N-k}}$.

from exerting effort level e_j relative to another effort level e_i is given by

$$\begin{aligned} & E\left[\int_{\tau_k}^{\tau_{k+1}} u_S(x_k(t,\tau_{[1,k]}))\,dt \,\middle|\, e_j\right] - E\left[\int_{\tau_k}^{\tau_{k+1}} u_S(x_k(t,\tau_{[1,k]}))\,dt \,\middle|\, e_i\right] \\ & = \int_{\mathbb{R}^k}\int_{\tau_k}^{\infty} f_k^{e_j}(\tau_{[1,k]})\ \mathrm{Prob}^{e_j}[\tau_{k+1}>t|\tau_k]\, e^{-\gamma t}\ u_S(x_k(t,\tau_{[1,k]})) \\ & \times \underbrace{\mathbf{P}_{k,e_i,e_j}(t;\tau_{[1,k]})}_{\text{1 net the relative likelihood of receiving } x_k(t,\tau_{[1,k]})} dt\, d\tau_{[1,k]}\,, \end{aligned} \tag{5}$$

where we have set[18]

$$\begin{aligned} \mathbf{P}_{k,e_i,e_j}(t;\tau_{[1,k]}) \ &\equiv\ 1 \ - \ \frac{\mathrm{Prob}^{e_i}[\tau_{k+1}>t|\tau_k]\, f_k^{e_i}(\tau_1,\cdots,\tau_k)}{\mathrm{Prob}^{e_j}[\tau_{k+1}>t|\tau_k]\, f_k^{e_j}(\tau_1,\cdots,\tau_k)} \\ &=\ 1 \ - \ \frac{p_{e_i}(\tau_1)\cdots p_{e_i}(\tau_k)(G_{e_i}(t))^{N-k}}{p_{e_j}(\tau_1)\cdots p_{e_j}(\tau_k)(G_{e_j}(t))^{N-k}}. \end{aligned} \tag{6}$$

The quantity $\mathbf{P}_{k,e_i,e_j}$ will play a fundamental role in the structure of the optimal contract. From the intermediary's point of view, $\mathbf{P}_{k,e_i,e_j}$ determines the likelihood of receiving the cash flows $x_k(t,\tau_{[1,k]})$ at time t under the effort e_j, relative to that under the effort e_i. From the investor's point of view, $\mathbf{P}_{k,e_i,e_j}$ is equal to 1 minus the likelihood of the event that the intermediary exerted effort e_i relative to the effort level e_j, given that the first k defaults occur at times $\tau_1,\cdots,\tau_k$ and the next default occurs no earlier than at time t. Formula (5) implies that the effect of the (IC) constraints on the structure of the optimal contract is completely determined by the nature of the functions $\mathbf{P}_{k,e_i,e_j}$.

Since both (IC) and (PC) will enter the first order conditions with the corresponding Lagrange multipliers $\{\mu_{\mathrm{IC}},\mu_{\mathrm{PC}}\}$, we will also need the following definition:

$$\Psi_k(t,\tau_{[1,k]};\{\mu_{\mathrm{IC}},\mu_{\mathrm{PC}}\}) \ \equiv\ \mu_{\mathrm{PC}} \ + \ \sum_{i\neq j}\mu_{\mathrm{IC},i}\mathbf{P}_{k,e_i,e_j}(t,\tau_{[1,k]})\,. \tag{7}$$

Finally, we will call a contract strictly incentive and participation constraint compatible

[18] See Appendix for a derivation of this expression.

if both (IC) and (PC) hold with strict inequalities. We can now describe the optimal contract implementing a given effort level j.

Theorem 3..1 *Fix an effort level e_j, $j > 1$ and suppose that the set of strictly incentive and participation constraint compatible contracts, implementing the effort level e_j is non-empty. Then, there exist Lagrange multipliers $\mu_{\text{PC}} \geq 0$, $\mu_{\text{IC},i} \geq 0$, $i \neq j$, such that the optimal contract $\{x_k(t, \tau_{[1,k]})$, $k = 0, \cdots, N\}$ satisfies*

$$\frac{e^{-rt} u_B'(\delta_k - x_k)}{e^{-\gamma t} u_S'(x_k)} = \Psi_k(t, \tau_{[1,k]}; \{\mu_{\text{IC}}, \mu_{\text{PC}}\}) \tag{8}$$

if

$$\Psi_k(t, \tau_{[1,k]}; \{\mu_{\text{IC}}, \mu_{\text{PC}}\}) \geq \frac{e^{-rt} u_B'(\delta_k)}{e^{-\gamma t} u_S'(0)} \tag{9}$$

and $x_k = 0$ otherwise.

Furthermore, the contract has a finite maturity: There exists a $\bar{T} > 0$ such that $x_k(t, \tau_{[1,k]}) = 0$ for all $t \geq \bar{T}$ and all $k \geq 0$.

The nature of the optimal contract is determined by the interaction of three forces: (i) (IC) constraints, driven by the dynamics of $\mathbf{P}_{k,e_i,e_j}$; (ii) benefits of paying the intermediary early, driven by the impatience wedge ($\gamma > r$); and (iii) risk sharing.[19] Intuitively, the contract incentivizes the intermediary by decreasing payments proportionally to the likelihood $1 - \mathbf{P}_{k,e_i,e_j}(t, \tau_{[1,k]})$ of a deviation from the desired level of effort e_j to an alternative effort e_i. The strength of these incentives is characterized by the size of the corresponding Lagrange multiplier $\mu_{\text{IC},i}$. At the optimum, the ratio of the discounted marginal loss for the investor, $e^{-rt} u_B'(\delta_k - x_k)$, and the discounted marginal benefit for the intermediary, $e^{-\gamma t} u_S'(x_k)$, in any given state is proportional to a linear combination

[19] Note that, without moral hazard and the impatience wedge (i.e., $\mu_{IC,i} = 0$ and $\gamma = r$), perfect risk-sharing is achieved.

of these likelihoods, given by Ψ_k, provided that the signal of no deviation to any alternatives of the seller implied by this state is sufficiently strong. If the signal of no deviation is not strong enough, (i.e., when (9) is violated), (LL) constraint binds and the optimal payments are reduced to zero.

To gain a deeper understanding of the optimal dynamic incentive provision mechanism, we need to investigate properties of the functions $\mathbf{P}_{k,e_i,e_j}(t,\tau_{[1,k]})$. Recall that the quantity

$$h_{e_j}(t) \;=\; \frac{p_{e_j}(t)}{G_{e_j}(t)}\,. \tag{10}$$

is commonly referred to as the default hazard rate. We need the following definition:

Definition 3..2 *We say that an effort level e_j leads to a lower default hazard rate than the effort level e_i if $h_{e_j}(t) \leq h_{e_i}(t)$ for all $t \geq 0$.*[20] *In this case we say that p_{e_j} dominates p_{e_i} in the hazard rate order, and we write $p_{e_i} \preceq_{hr} p_{e_j}$. Furthermore, we write $p_{e_i} \prec_{hr} p_{e_j}$ if the strict inequality $h_{e_j}(t) < h_{e_i}(t)$ holds for all t. Finally, we say that p_{e_j} dominates p_{e_i} in the likelihood ratio order if $p_{e_j}(t)/p_{e_i}(t)$ is monotone increasing in t. In this case we write $p_{e_i} \prec_{lr} p_{e_j}$.*

We can now state the following result.

Proposition 3..3 *The following is true*

- *If $p_{e_i} \preceq_{hr} p_{e_j}$, then $\mathbf{P}_{k,e_i,e_j}(t,\tau_{[1,k]})$ is monotone decreasing with k and hence, so is $x_k(t,\tau_{[1,k]})$ if $p_{e_i} \preceq_{hr} p_{e_j}$ for all $i \neq j$.*

- *If $p_{e_i} \prec_{lr} p_{e_j}$ ($p_{e_j} \prec_{lr} p_{e_i}$), then $\mathbf{P}_{k,e_i,e_j}(t,\tau_{[1,k]})$ is monotone increasing (decreasing) in default times $\tau_{[1,n]}$ and hence, so is $x_n(t,\tau_{[1,n]})$ if $p_{e_i} \prec_{lr} p_{e_j}$ ($p_{e_j} \prec_{lr} p_{e_i}$) for all $i \neq j$.*

[20] If a default time distribution has a lower hazard rate, it clearly also has a lower default risk.

The results of Proposition 3..3 are very intuitive. Indeed, if a default time distribution has a lower hazard rate, then for every time instant, the probability of default's happening instantaneously is smaller. Therefore, fewer defaults should occur over any given period of time. Hence, $\mathbf{P}_{k,e_i,e_j}(t,\tau_{[1,k]})$ is decreasing with k and, consequently, it is optimal to punish the intermediary for every new default. The likelihood ratio property is also very intuitive: It means that the likelihood of the effort e_j relative to the effort e_i is monotone increasing with the time of default and hence, so is $\mathbf{P}_{k,e_i,e_j}(t,\tau_{[1,k]})$. In other words, the later the default takes place, the higher is the likelihood of the effort e_j. This property naturally makes punishing the intermediary for early defaults optimal. As an illustration, note that in the HPT (2011) model, we have $p_{e_H}(t) = \lambda_L\, e^{-\lambda_L t}$, $p_{e_L}(t) = \lambda_H\, e^{-\lambda_H t}$ with $\lambda_L < \lambda_H$, and therefore both hazard rate and likelihood ratio properties hold: We have $\lambda_L = h_{e_H}(t) < h_{e_L}(t) = \lambda_H$ and $p_{e_L}(t) \prec_{lr} p_{e_H}(t)$.

4. The risk neutral limit

It is important to note that the simple form of the optimal contract is based on the assumption that the investor is able to control the intermediary's consumption, or, equivalently, that the intermediary cannot privately save. The introduction of private savings may significantly alter the optimal contract when the risk aversion of the intermediary is not too small. Indeed, in this case, the intermediary may undo incentives using private savings to smooth consumption. However, introducing private savings makes the optimal contracting problem significantly more complicated and investigating this important problem is beyond the scope of this paper. For this reason, everywhere in the sequel we will assume that the risk aversion of both market participants is sufficiently small.[21] It turns out that, when the desired effort level implements the lowest default hazard rate

[21]This is a natural assumption in a constant absolute risk aversion (CARA) setting. For example, an absolute risk aversion of 10^{-9} corresponds to a relative risk aversion of one for an agent with a billion dollar capital.

and the risk aversion of the market participants is sufficiently small, the optimal contract exhibits *extreme punishment for defaults:* It makes positive transfers to the intermediary only until the time of the first default. Namely, the following is true:

Theorem 4..1 *Suppose that p_{e_j} has the lowest hazard rate and both agents have exponential (CARA) preferences*

$$u_S(x) \;=\; A_S^{-1}(1-e^{-A_S x})\;,\;\; u_B(x) \;=\; A_B^{-1}(1-e^{-A_B x})\,.$$

Then, if A_B, A_S are sufficiently small and A_B/A_S is not too large, the optimal contract $\{x_k\}$ implementing effort e_j has $x_k \equiv 0$ for all $k \geq 1$.

The intuition behind this result is based on Proposition 3..3: When the desired effort level leads to the lowest hazard rate, $\mathbf{P}_{k,e_i,e_j}$ is monotone decreasing with k and hence, so is $x_k(t, \tau_{[1,k]})$. Therefore, the optimal contract always makes the largest transfers to the intermediary in the period $[0, \tau_1]$ before the first default occurs. When agents are sufficiently close to being risk neutral, concentrating all payments in this time interval is optimal because the compensation that the intermediary requires for taking the risk of early defaults is sufficiently small.

To continue the analysis of the risk-neutral limit, we need to impose additional technical conditions.

Definition 4..2 *We say that default time distributions are k-regular if the function $e^{-(\gamma-r)t}\,\Psi_0(t;\{\mu_{\rm IC},\mu_{\rm PC}\})$, defined in (7), has at most k local maxima for $t \in [0,+\infty)$ for any Lagrange multipliers $\mu_{\rm PC} \geq 0$, $\mu_{{\rm IC},i} \geq 0,\; i \neq j$.*

The following is true:

Theorem 4..3 *Under the hypothesis of Theorem 4..1, suppose that default time distributions are k-regular. Then, if $A_B, A_S \to 0$ so that A_B/A_S stays bounded, any contract in the limit set is of the following form:*[22]

There exists a $\kappa \in \{0, \cdots, k-1\}$, an increasing sequence of time instants $0 \leq t_0 < \cdots < t_\kappa < \infty$[23] *and a sequence $y_i \in \mathbb{R}_+$, $i = 0, \cdots, \kappa$, such that the optimal schedule of transfers to the intermediary is given by*

$$\sum_{i=0}^{\kappa} \mathbf{1}_{t=t_i} \; \mathbf{1}_{t_i<\tau_1} \, y_i \, .$$

To understand the intuition behind the optimal contract of Theorem 4..3, note that when both agents are risk neutral, the most efficient way to provide incentives for a given effort level is to concentrate all the contract's payments in the time instants with the highest likelihood of the desired effort level[24]. When the effort level e_j leads to the lowest default hazard rate, these likelihood-maximizing time instants t_i, $i = 0, \cdots, \kappa$ always belong to the time proceeding the first default, as we show in Theorem 4..1; the number κ of these instants is bounded from above by k, the maximal number of local maxima of the weighted likelihood. The sizes y_i of the corresponding payments are determined by the severity of the corresponding IC constraints. This contract structure can also be well understood from the tradeoff between learning and impatience. Clearly, investors would like to postpone payments so that they can learn more from observing a longer history of defaults. However, delaying payments is costly because of the relative impatience of the

[22]Formally, it means that, for any sequence of risk aversions converging to zero, we can pick a subsequence such that the corresponding contracts converge to a contract of the from described in the theorem. If there is a unique optimal contract in the risk neutral limit, then the convergence takes place in the standard sense.

[23]It follows from the proof in the Appendix that there exist Lagrange multipliers $\mu_{\text{PC}} \geq 0$, $\mu_{\text{IC},i} \geq 0$, $i \neq j$ such that t_i is a local maximum of $e^{-(\gamma-r)t}\Psi_0(t; \{\mu_{\text{IC}}, \mu_{\text{PC}}\})$ and $e^{-(\gamma-r)t}\Psi_0(t_i; \{\mu_{\text{IC}}, \mu_{\text{PC}}\}) = 1$ for all $i = 0, \cdots, \kappa$. This observation can be used directly to determine t_i.

[24]With multiple effort levels, the "highest likelihood of the desired effort" should be interpreted as "highest $\sum_{i \neq j} \mu_{\text{IC},i} \mathbf{P}_{k,e_i,e_j}(t, \tau_{[1,k]})$."

intermediary. This leads to an optimal incentive provision mechanism characterized by *multiple waiting periods:* The intermediary gets partially remunerated for the absence of defaults during every subsequent waiting period $[t_i, t_{i+1})$. So, if no default occurs until $t = t_0$, the intermediary receives the first payment y_0 and enters a waiting phase $(t_0, t_1]$. If no default occurs until $t = t_1$, the intermediary receives the next payment y_1, etc., until either the last payment gets paid at time t_κ, or a default occurs, in which case the intermediary loses all subsequent payments.

It is important to relate our main results (Theorems 4..1 and 4..3) to those of HPT (2011). Namely, HPT (2011) obtains a special case of Theorem 4..1, assuming from the beginning that (1) both agents are risk neutral (without taking the limit $A_S, A_B \to 0$); (2) the effort choice is binary; and (3) default time distributions are exponential. In this case, it is possible to show that $p_{e_L} \prec_{hr} p_{e_H}$, and the default distributions are 1-regular, so that the optimal contract is characterized by a single payment after a single waiting period. Being the first of its kind, the result of HPT (2011) is important for understanding the nature of optimal incentive provisions for securitization. In particular, HPT (2011) identify two key economic forces that determine the shape of the optimal contract: (1) the investor wants to pay the intermediary as soon as possible to exploit the impatience wedge and (2) an interaction between the limited liability and incentive compatibility constraints creates value for information "quality". Information quality improves over time and therefore the *timing* of payments is a major incentive provision mechanism.

Theorems 4..1 and 4..3 show that the same economic intuition still holds in a more general setting and the results of HPT (2011) are robust. Namely, Theorem 4..1 shows that extreme punishment for defaults, a property of the optimal contract that HPT (2011) obtained in the risk neutral setting, is robust to small perturbations to risk neutrality so long as higher effort leads to lower hazard rates. Theorem 4..3 shows that a

lumpy contract is optimal in the risk neutral limit, as in HPT (2011); the only difference is that multiple payment dates may arise. For example, when default distributions are exponential, it is possible to show that they are $(K-1)$-regular, with K being the number of possible effort levels. Therefore, the following is true.

Proposition 4..4 *Suppose that* $p_{e_i}(t) = \lambda_i\, e^{-\lambda_i\, t}$ *for all* $i = 1, \cdots, K$, *with* $\lambda_1 > \cdots > \lambda_K$. *Then, default time distributions are* $(K-1)$*-regular, and therefore the result of Theorem 4..3 holds with* $k = K-1$.

The fact that multiple waiting periods are optimal in the multiple effort case is quite intuitive: Each payment date prevents the intermediary from deviating to the corresponding alternative effort level. Namely, the shortest waiting period $(0, t_0)$ incentivizes the intermediary to exert an effort higher than e_1, the second payment (at time t_1) provides incentives to exert an effort higher than e_2, etc.

Theorem 4..3 can be used to explicitly calculate the optimal payments y_i, as well as the optimal payment times t_i, for any number of possible effort levels. However, for simplicity, everywhere in this section we confine ourselves to the binary effort case. Furthermore, we will often assume that the default time distributions come from the Black and Cox (1976) structural default model.[25] Namely, we assume that

$$p_{e_j}(t) \;=\; \frac{a_j}{\sqrt{2\pi}\,\sigma_j\, t^{3/2}}\, e^{-\frac{(m_j t + a_j)^2}{2\sigma_j^2 t}} \;,\; j = H,\; L \tag{11}$$

for some $a_j, m_j, \sigma_j > 0$. In this case, it is possible to show[26] that $p_{e_L} \prec_{hr} p_{e_H}$ if and only if $\frac{a_H}{\sigma_H} > \frac{a_L}{\sigma_L}$ and $\frac{m_H a_H}{\sigma_H^2} > \frac{m_L a_L}{\sigma_L^2}$. Define

$$\phi_1(t) \;\equiv\; \frac{(e^{(\gamma-r)t} - 1)}{1 - (G_{e_L}(t)/G_{e_H}(t))^N} \text{ and } t_1^* \;\equiv\; \arg\min_{t \geq 0}\; \phi_1(t)\,. \tag{12}$$

[25] See Section 6. for a detailed discussion of this model.

[26] See Proposition B.2 in the Appendix.

The following is true:

Theorem 4..5 *Suppose that the effort is binary, default time distributions are from the Black and Cox model,* $p_{e_L} \prec_{hr} p_{e_H}$*, and that the desired effort level is* e_H*. Then, in the risk-neutral limit, the optimal contract makes a lump sum payment* $y_0 \geq 0$ *to the intermediary at time* 0*, and then a lump sum payment* $y_1 > 0$ *at a time* $t^* > 0$ *if no default occurs before* $t = t^*$*. Furthermore, there exists a threshold* C^* *such that for* $C_H < C^*$*, we have:*

$$t^* \;=\; t_1^* \;,\; y_1 \;=\; \frac{e^{\gamma t_1^*}\,(C_H - C_L)}{(G_{e_H}(t_1^*))^N - (G_{e_L}(t_1^*))^N}\,,$$

and $y_0 > 0$.

Figure 1 provides an illustration for the convergence result of Theorem 4..5.

Insert Figure 1 Here

A very important consequence of Theorem 4..5 is the optimality of a strictly positive payment to the intermediary at time zero when the cost of effort is not too high.[27] To understand the intuition behind this result, recall that the structure of the optimal contract is determined by optimal tradeoff between waiting for better information (in order to reduce the cost of incentive provision) and paying the intermediary early in order to exploit the impatience wedge between the intermediary and the investor. This tradeoff is determined by the dynamics of the functions $\mathbf{P}_{k,e_L,e_H}$ and $e^{-(\gamma-r)t}\Psi_k$, representing the "discounted value of incentives". By Theorem 4..1, extreme punishment for defaults is

[27]Our numerical results indicate that the condition $C_H < C^*$ of Theorem 4..5 is absolutely non-restrictive. In fact, the threshold C^* is so high (higher than the total value of the assets in the pool) that the inequality $C_H < C^*$ holds for any reasonable parameter values. For the sake of completeness, the case of very high effort cost (i.e., $C_H > C^*$) is considered in the Appendix.

optimal and we can confine ourselves to the case $k = 0$, with

$$e^{-(\gamma-r)t}\,\Psi_0(t) \;=\; e^{-(\gamma-r)t}\,(\mu_{\rm PC} + \mu_{\rm IC}\,\mathbf{P}_{0,e_L,e_H}(t))\;,$$

and

$$\mathbf{P}_{0,e_L,e_H}(t) \;=\; 1 \;-\; \left(\frac{G_{e_L}(t)}{G_{e_H}(t)}\right)^N\;.$$

If higher effort reduces default hazard rate, the function $\mathbf{P}_{0,e_L,e_H}(t)$ is monotone increasing with t. The derivative

$$\mathcal{S}(t) \;\equiv\; \frac{d}{dt}\mathbf{P}_{0,e_L,e_H}(t) \tag{13}$$

can therefore be interpreted as the speed of information arrival. If the investor and the intermediary are equally patient, it is always optimal to wait for more information and indefinitely postpone payments to the intermediary. However, when $\gamma > r$, waiting is costly and the interplay between the size $\gamma - r$ of the impatience wedge and the speed of information arrival determines the optimal timing of payments to the intermediary. Clearly, when the effort cost C_H is sufficiently high, incentive provision is very costly and only a very large delayed payment can satisfy the (IC) constraint. The (PC) constraint is then satisfied automatically. In contrast, when C_H is sufficiently small, the moral hazard problem is mild and even a small delayed payment is sufficient to provide the necessary incentives. In this case, a positive initial payment y_0 may be the optimal way to satisfy the (PC) of the intermediary. One very intuitive sufficient condition is that the speed of information revelation is zero at time zero. Indeed, in this case it may take a long time for better information to arrive and paying the intermediary immediately is optimal. This intuition is formalized in the following proposition

Proposition 4..6 *Suppose that higher effort reduces default hazard rate, the densities $p_{e_H}(t), p_{e_L}(t)$ are continuous for small t, and the optimal contract is of the form described*

in Theorem 4..5. If

$$\mathcal{S}(0) \;=\; 0\,,$$

then there exists a threshold C^* such that optimal the payment y_0 is strictly positive for all $C_H < C^*$.

As an illustration of this result, let us consider the following modification of the HPT (2011) model: Suppose that there exists an $s > 0$ such that no default occurs until time s.[28] That is,

$$p_{e_H} \;=\; \mathbf{1}_{t\geq s}\,\lambda_L\,e^{-\lambda_L(t-s)}\;,\; p_{e_L} \;=\; \mathbf{1}_{t\geq s}\,\lambda_H\,e^{-\lambda_H(t-s)} \tag{14}$$

for some $\lambda_H > \lambda_L > 0$. Then, it is possible to show that the optimal contract for this model has the form, characterized in Theorem 4..5. However, since no information is revealed up to time s, we ought to have $t^* > s$. When the delay s is sufficiently large, the impatience wedge makes a positive initial payment optimal. The following is true.

Proposition 4..7 *Suppose that the default time distributions are given by (14). Then, the optimal contract has the form, described in Theorem 4..5 with some $y_0(s), y_1(s), t^*(s)$. Furthermore, if the cost of effort satisfies*

$$C_H \;<\; C_L \;+\; (U_S^0 + C_L)\,\frac{(\lambda_H - \lambda_L)N}{\gamma - r}\,, \tag{15}$$

then, there exists a $\bar{s} > 0$ such that $y_0(s) > 0$ if and only if $s > \bar{s}$. By contrast, if (15) is violated then $y_0 = 0$ for all $s > 0$.

When the speed of information revelation is non-zero at time zero, the situation is more subtle. What matters then is not only the speed $\mathcal{S}(t)$ (see (13)) of information arrival, but also the rate at which $\mathcal{S}(t)$ is changing. Suppose that the speed of information

[28] We thank an anonymous referee for suggesting this very nice example.

arrival is monotone increasing with time. Then, if the rate at which the speed is increasing is sufficiently large (larger than the impatience wedge $\gamma - r$), it is optimal to delay the payments further and wait until even "better" information starts arriving. In this case, it may be optimal to make a small payment to the intermediary at time zero and then wait until the speed starts increasing really fast, and only then make the next payment. The following is true.

Proposition 4..8 *Suppose that higher effort reduces default hazard rate, the densities $p_{e_H}(t), p_{e_L}(t)$ are continuously differentiable for small t, and the optimal contract is of the form, described in Theorem 4..5. If $\mathcal{S}(0) > 0$ and*

$$\frac{\frac{d}{dt}\mathcal{S}(t)}{\mathcal{S}(t)}\Big|_{t=0} > \gamma - r$$

then there exists a threshold $C^* > 0$ such that the optimal payment y_0 is strictly positive if and only if $C_H < C^*$.

As an example, consider the HPT (2011) model. Then,

$$\frac{d}{dt}\mathcal{S}(t) = -((\lambda_H - \lambda_L)N)^2\, e^{-(\lambda_H - \lambda_L)Nt} < 0,$$

and therefore the speed of information arrival is decreasing over time and Proposition 4..8 is not applicable. As HPT (2011) show, y_0 is always equal to zero in this case.

The following proposition provides some useful comparative statics results about the maturity t^* of the optimal contract and the optimal payment y_1.

Proposition 4..9 *Under the hypothesis of Theorem 4..5, the maturity $t^* = t_1^*$ of the optimal contract is always monotone decreasing in N and $\gamma - r$ and is increasing in the size of default risk under high effort. The payment y_1 is increasing in $\gamma - r$, and decreasing*

in N and in the size of default risk under high effort. Furthermore, t^* converges to 0 as $N \to \infty$.

Intuitively, maturity t^* and the size y_1 of the delayed payment are determined by the severity of the moral hazard problem and the conflict of interest between the investors and the intermediary. If exerting high effort only marginally decreases default risk, the moral hazard problem is severe and the optimal waiting period is long. When the intermediary is relatively impatient, delaying payments is costly for the investor, making it optimal to reduce the waiting period $[0, t^*)$. To compensate for this cost and to improve incentives, delayed payment y_1 has to be made larger. Finally, the fact that the maturity t^* is decreasing with N is justified by the information enhancement effect of pooling, emphasized in HPT (2011): When N is large, investors can learn much faster about the intermediary's effort, simplifying the problem of optimal incentive provision.

The elegant characterization of the optimal contract in the risk-neutral limit, provided in Theorem 4..3, critically depends on the assumption that the desired effort level implements the lowest default hazard rate. When this assumption is violated, Theorem 4..1 does not hold in general and extreme punishment for defaults might not be optimal. Consequently, in the risk-neutral limit, the optimal contract might make positive payments to the intermediary even after a few defaults have occurred. In our benchmark model with binary effort and Black and Cox default time distributions, Proposition B.2 implies that high effort does not lead to a lower hazard rate if and only if $m_H/\sigma_H < m_L/\sigma_L$. In this case, for a large t, the default hazard rate $h_{e_L}(t)$ under low effort is significantly lower than the rate given a high effort, and the results of Theorems 4..1 and 4..3 do not hold in general. The following is true:

Theorem 4..10 *Suppose that the effort is binary, default time distributions are generated from the Black and Cox model, and $m_H/\sigma_H < m_L/\sigma_L$. Suppose also that $\gamma - r$ is sufficiently small. Then:*

- *Conditional on observing n defaults, maximal transfers to the intermediary happen when these defaults take place immediately one after another, and we have*

$$\varphi_k(t) \;\equiv\; \max_{\tau_{[1,k]}\leq t} x_k(t;\tau_{[1,k]}) \;=\; x_k(t;(t,t,\cdots,t))\,;$$

- *The maximum of $\varphi_k(t)$ over t is monotone increasing in the number n of defaults; and*

- *There exist $n^*>0$ and $t^*>0$ such that, in the limit as $A_S, A_B \to 0$ so that A_B/A_S stays bounded, the contract only makes transfers to the intermediary when n^* defaults occur at time instants, sufficiently close to τ^*, and with one immediately after another.*

As we have explained, in the case when both agents are risk neutral, the optimal incentive alignment mechanism can be implemented by concentrating payments in the state that maximizes the likelihood of a high effort level; however, the timing of transfers has to be adjusted for the relative impatience of the intermediary ($\gamma > r$). Because the desired effort level does not implement the minimal hazard rate, $\mathbf{P}_{k,e_L,e_H}(t,\tau_{[1,k]})$ increases with the number k of defaults when t is sufficiently large. When the intermediary is sufficiently impatient, postponing payments into the future is too costly for the investor, and the optimal contract will take the same form as in Theorem 4..1. However, when γ is sufficiently close to r, waiting for a few defaults to occur (namely, n^* defaults in Theorem 4..10) is optimal, and the contract takes the form described. We refer to this type of payment as *paying for default cascades.*

It should be pointed out that the optimal contract, described in Theorem 4..10 has a very unattractive feature of being highly vulnerable to default manipulation by the intermediary. Indeed, if the payment is only made to the intermediary following a cascade of defaults, the intermediary has strong incentives to collude with the borrowers

and arrange that they default at the "optimal" time. This sensitivity to intermediary manipulation is more troubling than that to investor manipulation since intermediaries typically have direct access to borrowers.[29]

5. Securitization and the optimal effort level

In this section, we use our explicit expressions for the optimal contract in the risk-neutral case to study the effect of securitization on the equilibrium effort choice. For simplicity, we assume that the effort is binary (i.e., $e \in \{e_H, e_L\}$), and the effort costs are proportional to the number of assets in the pool (i.e., $C_j = N\, c_j\ ,\ j = H, L$ for some $c_L,\ c_H > 0$). The outside option of the intermediary is simply retaining the initially created asset pool (the "originate-and-retain" model), that is,

$$U_S^0 \ = \ \max_j U_S(\{\delta_n\}, e_j)\,,$$

and the outside option of the investor is zero: $U_B^0 = 0$.

In the absence of moral hazard, the agent who has the bargaining power can extract full surplus from the counterparty. In our case, this *first best* surplus (conditional on a given effort level e_j) is equal to

$$\mathrm{FB}_j(N) \ = \ U_B(\{\delta_n\}, e_j) - (U_S^0 + C_j) \ = \ N \cdot \mathrm{FB}_j(1) \tag{16}$$

and is proportional to the number of assets in the pool.

We start with the following simple observation: Because the intermediary is relatively impatient ($\gamma > r$), providing incentives is costly. Therefore, the most efficient way to implement low effort is simply to make a single payment at time zero. This is formalized

[29] We thank an anonymous referee for this important observation.

in the following proposition.

Proposition 5..1 *In the risk-neutral case, the optimal contract for implementing low effort is to pay the intermediary a fixed lump sum at time zero. Furthermore, we have:*

- *In the competitive case, this lump sum equals* $U_B(\{\delta_n\}, e_L)$ *(full surplus extraction by the intermediary).*
- *In the monopolistic case, the lump sum equals* $U_S^0 + C_L$ *(full surplus extraction by the investor).*

In particular, total surplus coincides with FB_L.

In the high effort case, (IC) constraint always limits the set of contracts and reduces total surplus.[30] HPT (2011) show that pooling has an *information enhancement effect:* Increasing the number of securitized assets simplifies the problem of incentive provision and makes it less costly. The following proposition shows that, in fact, the total surplus loss vanishes in the limit when the number of securitized assets becomes large.

Proposition 5..2 *Under the hypothesis of Theorem 4..5, the total second best surplus* SB_H *for high effort is independent of bargaining power allocation and is given by:*

$$\mathrm{SB}_H(N) \;=\; N\left(\mathrm{FB}_H(1) - (c_H - c_L)\,\phi_1(t_1^*)\right). \tag{17}$$

The total surplus loss per asset $(\mathrm{FB}_H(N) - \mathrm{SB}_H(N))/N$ *is monotone decreasing in the number* N *of assets and converges to zero as* $N \to \infty$.

Now we are ready to discuss the optimal effort choice in the presence of securitization. Because $\gamma > r$, the increase in the pool value resulting from a higher screening effort is

[30] IC constraint always binds for the optimal contract implementing high effort. Indeed, if it does not bind and the investor is risk neutral, the optimal contract makes deterministic payments independent of the default history, in which case the intermediary optimally chooses a low effort.

always higher for the investor than for the intermediary; that is,

$$U_B(\{\delta_n\}, e_H) \; - \; U_B(\{\delta_n\}, e_L) \; > \; U_S(\{\delta_n\}, e_H) \; - \; U_S(\{\delta_n\}, e_L)\,. \tag{18}$$

By (16), the optimal effort level is high in the first best case if and only if

$$U_B(\{\delta_n\}, e_H) \; - \; U_B(\{\delta_n\}, e_L) \; > \; C_H - C_L\,. \tag{19}$$

Indeed, if the cost of effort is higher than the increase in the market value of the pool, it will never be optimal for market participants to choose high effort. Because the second best surplus is always lower than the first best one, condition (19) is also necessary for high effort to be optimal in the second best case. Combining this observation with Proposition 5..2, we arrive at the following result.

Proposition 5..3 *In the presence of securitization, the equilibrium effort level is e_H if and only if the following is true:*

$$U_B(\{\delta_n\}, e_H) \; - \; U_B(\{\delta_n\}, e_L) \; \geq \; (1 + \phi_1(t_1^*))(C_H - C_L)\,. \tag{20}$$

Consequently:

- *If (19) does not hold, then the equilibrium effort level is e_L, both with and without securitization and independent of bargaining power allocation.*
- *If (19) holds, then there exists an $N^* \geq 1$ such that the equilibrium effort level is e_H if and only if $N \geq N^*$. In particular, for sufficiently large N, securitization always improves the equilibrium screening effort.*

Proposition 5..3 has direct implications for regulating securitization. Imposing a lower bound on the number of assets in the securitized pool and using the correct incentive

provision mechanisms might significantly improve an intermediary's screening incentives. To provide a better understanding of this effect, we state the following proposition.

Proposition 5..4 *Consider a* 1*-parameter family of distributions* $G(t,\alpha)$*, such that* $G(t,\alpha)$ *is continuous and increases in* α *in the hazard rate order. Suppose that* $G_{e_j}(t) = G(t,\alpha_j)$*,* $j = H, L$*. Fix the parameter* α_L*. Then, there exist thresholds* $\overline{\alpha} > \underline{\alpha}$ *in* $(\alpha_L, +\infty)$ *such that:*

- *Without securitization, the intermediary chooses high effort if and only if* $\alpha_H > \bar{\alpha}$*; and*
- *In the first best case, the investor chooses a contract implementing high effort of the intermediary if and only if* $\alpha_H > \underline{\alpha}$*; and*
- *For all* $\alpha_H \in (\underline{\alpha}, \overline{\alpha})$*, there exists a threshold* $N^*(\alpha_H)$ *such that the equilibrium effort level is high if and only if* $N > N^*(a_H)$*. Consequently, for all* $\alpha_H \in (\underline{\alpha}, \overline{\alpha})$*, securitization strictly improves the equilibrium screening effort if and only if* $N > N^*(\alpha_H)$*.*

Figure 2 illustrates how the minimal number N^* of assets in the pool depends on various model parameters. For example, we can see that, even when screening increases the distance to default by 30% (from $a_L = 0.2$ to $a_H = 0.27$), the intermediary will choose a low effort without securitization, whereas, with just a few hundred assets in the pool, securitization makes a high effort optimal even for $a_H = 0.23$.

Insert Figure 2 Here

6. The Black and Cox Default Time Distributions

In this section, we provide a detailed analysis of a special case of our general model in which the default times are generated by the Black and Cox (1976) structural default

model.[31] In this model, the borrower defaults when a given stochastic process X_t falls below a given threshold X_B. This process can be interpreted as the operational cash flows (when the borrower is a firm) or income (when the borrower is an individual), or the house price (for the case of a mortgage). The process X_t is assumed to follow a geometric Brownian motion, with the drift μ and volatility σ,

$$dX_t \;=\; X_t\,(\mu\,dt \,+\, \sigma\,dB_t)\;,$$

where B_t is a standard Brownian motion. We assume that X_t is non-observable for the outside investors. This is definitely true in the case of consumer loans or mortgages, and it is usually justified in the literature by incomplete accounting information for the case when the borrower is a firm (Duffie and Lando, 2001). Although the intermediary might have some information about X_t for $t > 0$, we assume for simplicity that this information is non-contractible and that the contractual payments can only be conditioned on the default history.[32] The probability distribution of the default time τ^{X_B}, the first time X_t falls below X_B, is given by

$$\text{Prob}[\tau^{X_B} < t] \;=\; F^{a,m,\sigma}(t) \;\equiv\; 1 \;-\; \Phi\left(\frac{mt+a}{\sigma\sqrt{t}}\right) \;+\; e^{\frac{-2ma}{\sigma^2}}\,\Phi\left(\frac{mt-a}{\sigma\sqrt{t}}\right)\;, \qquad (21)$$

which has the density

$$p^{a,m,\sigma}(t) \;=\; \frac{a}{\sqrt{2\pi}\,\sigma\,t^{3/2}}\,e^{-\frac{(mt+a)^2}{2\sigma^2 t}} \qquad (22)$$

with

$$m \;= \mu - 0.5\sigma^2\;,\; a \;=\; \log(X_0/X_B) \;>\; 0\,.$$

[31] The Black and Cox (1976) model has now become a benchmark for calculating default probabilities. See Duffie and Singleton (2003) for a detailed analysis and applications.

[32] For example, this information may be soft and difficult to transmit to outside investors. Alternatively, if the number N of assets in the pool is large, monitoring all of them is extremely costly for outside investors.

Here,

$$\Phi(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$$

is the cumulative distribution function (CDF) of a standard normal distribution. For simplicity, we always assume that the (risk adjusted) growth rate m is positive for all borrowers (i.e., $m > 0$). A very important property of Black and Cox default time distributions it that they are *defective:* There is a strictly positive probability $P_{\infty}^{a,m,\sigma}$ that the borrower will never default. Namely,

$$\text{Prob}[\tau^{X_B} = \infty] = P_{\infty}^{a,m,\sigma} \equiv 1 - e^{-\frac{2ma}{\sigma^2}}. \tag{23}$$

This defectiveness is a very important economic phenomenon that appears in any model in which there is growth: When the stochastic distance to default, $\log(X_t/X_B)$, is growing with positive probability, the borrower might never default. As we will show, this property is responsible for several surprising features of the optimal contract.

The following proposition characterizes the dependence of $p^{a,m,\sigma}$ on the parameters (a, m, σ).

Proposition 6..1 *The density $p^{a,m,\sigma}$ is*

(1) increasing in a, m and decreasing in σ in the sense of $\prec_{hr}$ order;

(2) increasing in a in the sense of the $\prec_{lr}$ order;

(3) decreasing in m with respect to the $\prec_{lr}$ order; and

(4) neither increasing nor decreasing in σ with respect to the $\prec_{lr}$ order.

Property (1) is very intuitive: Clearly, a borrower with a higher initial distance to default a, a higher growth rate m, or a lower volatility σ will default with a lower

probability at every instant. Property (2) is also to be expected: Because borrowers that have higher initial capital default later, on average, the likelihood of the borrowers having higher initial capital should be monotone increasing in the time of default. However, property (3) is counterintuitive and surprising: It means that the later a default occurs, the higher is the likelihood that the borrower has a *low* cash flow growth rate. The reason is that the density $p^{a,m,\sigma}$ is defective. A borrower who has a higher m, also has a higher probability of never defaulting. However, conditional on the event that a default occurs in finite time, it happens earlier, on average, for a borrower with a higher growth rate m.

In the Black and Cox setting described above, the intermediary's screening effort has a very clear meaning: He can screen the borrowers for their initial distance to default a, their growth rate μ, and their volatility σ. We assume that, conditional on an effort level e_j, all borrowers in the pool have the same parameters (a_j, m_j, σ_j).[33] Although the actual screening procedure implemented by banks is more complicated and relies to a significant degree on soft information that is difficult to quantify, the three parameters (a, m, σ) have a very clear economic meaning and can be directly related to observable quantities. The parameter a can be clearly associated with the borrower's initial creditworthiness and is therefore the easiest to estimate. The cash flow volatility σ can also be estimated if a sufficient amount of past income/cash flow/house price data is available. The growth rate parameter m is the most difficult one to estimate empirically. However, it is possible to use observable information to get some rough idea about its magnitude. For example, for consumer loans, information on education level and past employment and associated income growth can be used to provide information about μ. In the case of commercial loans for a real estate company, rental income and length of leases could be used to

[33]Alternatively, we might assume that the parameters (a_j, m_j, σ_j) are themselves random and are sampled from a probability distribution, determined by the effort level. Our analysis can be easily extended to this more general setting.

estimate μ. However, even if a higher screening effort decreases default risk, it does not necessarily lead to an increase in both a, m and a decrease in the volatility σ. The following proposition describes the effect of screening potential borrowers on the basis of a, m and σ on the default time distribution.

Combining Proposition 6..1 with Proposition 3..3, we get that the optimal contract might exhibit several surprising features:

(1) If the intermediary is able to select borrowers with higher growth, then the optimal payments are decreasing with the number of defaults. However, when defaults occur after a sufficiently long time period, the optimal contract exhibits *punishment for late defaults:* The intermediary will be payed less if the defaults occur too late;

(2) If the intermediary selects borrowers with high initial distance to default but with low growth, the optimal contract will exhibit *punishment for too few defaults:* Payments to the intermediary will *increase* with the number of defaults when these defaults occur sufficiently late. Nevertheless, after a fixed number n of defaults is observed, the intermediary will be paid more if these defaults occurred late.

Indeed, selecting borrowers with a higher growth rate increases the probability that the defaults will never occur. However, if the defaults do occur, they should happen early. Therefore, contracts that punish for late defaults is an efficient way to provide correct incentives. The same intuition applies to item (2): If the cohort of borrowers selected with the desired effort level is supposed to have lower growth rates, the corresponding default hazard rates would be higher for sufficiently large t. Therefore, defaults will tend to be more concentrated, and the optimal contract punishes the intermediary for too few defaults.

7. Conclusion

The recent turmoil in the asset-backed securitization market has led regulators and market participants to reflect upon the role of misaligned incentives in disrupting the efficient functioning of both financial and real sectors of the economy. Given the important benefits of securitization, including better risk sharing and a reduced cost of capital for the intermediary (see, e.g., Pennacchi, 1988), a properly organized securitization market might significantly improve society's welfare. However, natural asymmetric information and moral hazard problems between intermediaries and investors can lead to serious dysfunctionality and illiquidity in these markets. The collapse of highly rated structured finance products in 2007-2008 has obviously shown that credit ratings failed to resolve these important problems. Therefore, new mechanisms need to be developed, providing intermediaries with better incentives for monitoring the credit risk of securitized assets. Although new regulatory requirements (e.g., 5% retention rule) introduced in the Dodd-Frank Wall Street Reform and Consumer Protection Act are aimed at improving intermediaries' incentives, their efficiency may be highly sensitive to a particular economic environment.

In this paper, we study how efficient incentive alignment mechanisms can be designed to (partially) resolve the problem of incentive alignment in the securitization market. We show that the structure of optimal securitization contracts depends in a very non-trivial way on the nature of the underlying credit risk. In stark contrast to the conventional wisdom, we find that securitization *improves the intermediary's monitoring incentives,* even when the intermediary has full bargaining power in designing the optimal contract. The reason is that, when investors are sufficiently sophisticated (fully rational), they correctly anticipate the intermediaries' monitoring effort and optimally respond to this level of effort by requiring a higher credit risk premium. This endogenous incentive provision mechanism naturally improves intermediaries' incentives and leads to more efficient risk

allocation. Despite its being abstract and theoretical, our model has many realistic features: It allows us to take into account of risk aversion of all market participants, as well as for various sources of the credit risk of the assets collateralizing a structured product. We believe that future research, introducing more realistic market features (e.g., regulatory constraints, macroeconomic risk, hedging and saving for both investors and intermediaries), may lead to contracts that can eventually become an industry standard. Incorporating systemic risk (e.g., risks of recessions or stochastic depreciation in home prices) is another important direction in which our results can be extended. For example, in addition to determining the average default risk of the securitized assets, intermediary's effort might also affect the amount of systemic risk in the securitized portfolio. This may potentially lead to systemic crises, as in Farhi and Tirole (2010), and one can study efficient ways of preventing these welfare-decreasing outcomes.

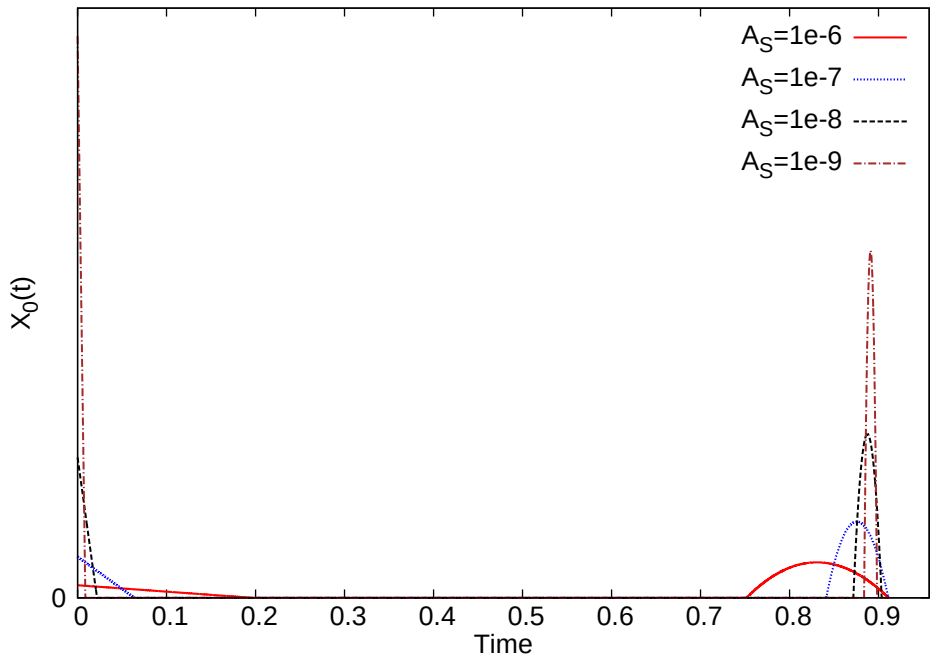
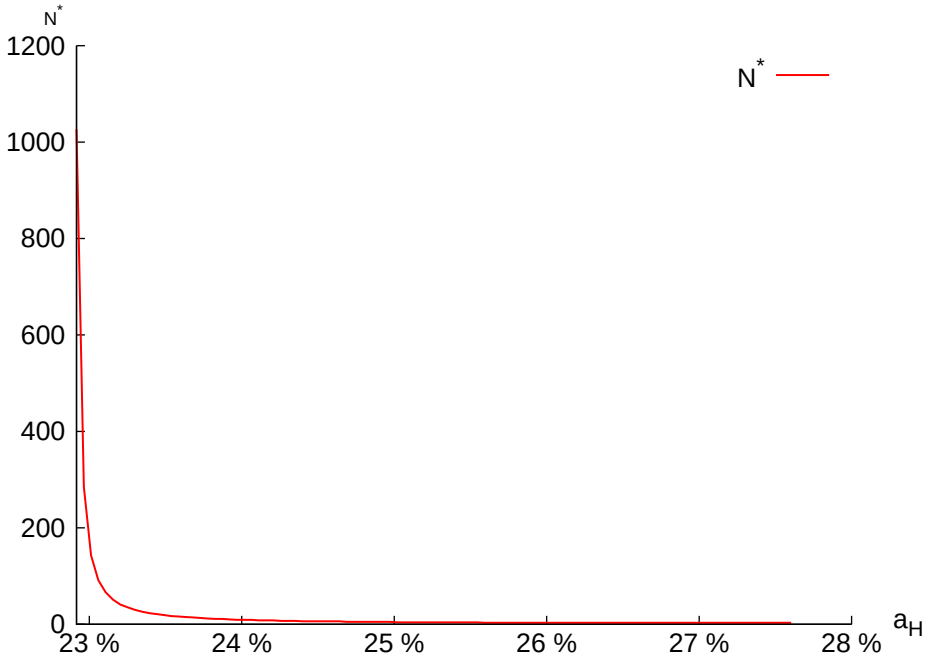
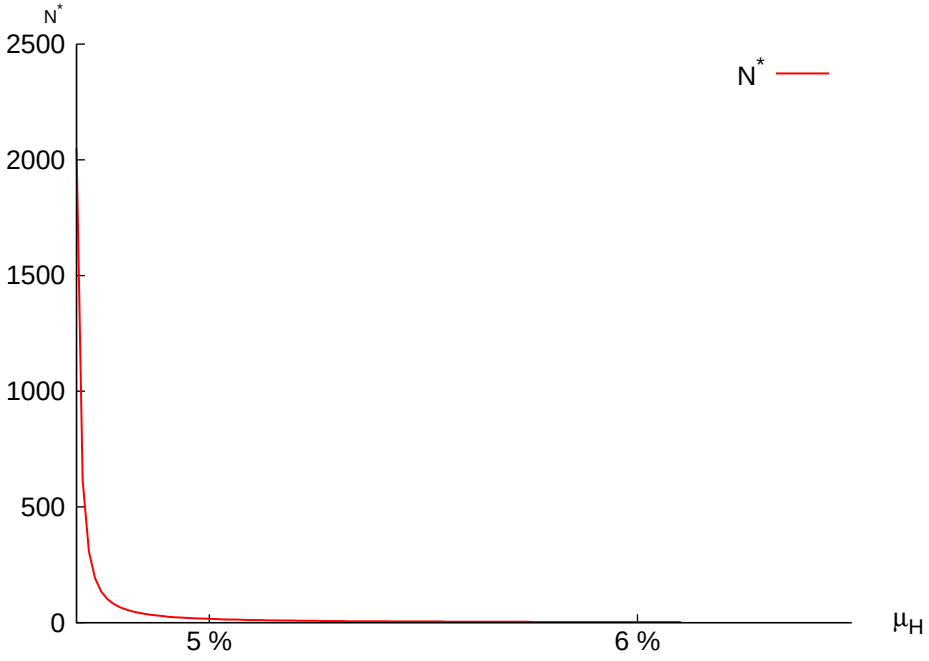


Figure 1: Convergence to the lump sum payments (Theorem 4..1) at $t_1^* \approx 11$ months as the intermediary's risk aversion A_S goes to zero. The unit of horizontal axis is one year. Parameter values: $A_B = 0,\ r = 5\%,\ \gamma = 10\%,\ C = (2\%)\frac{Nu}{r},\ N = 100,\ R/u = 50\%,\ u = 1,\ (a_H, \mu_H, \sigma_H) = (0.4, 0.04, 0.1),\ (a_L, \mu_L, \sigma_L) = (0.2, 0.04, 0.1)$ (Growth rate and volatility are taken on a yearly basis.)



(a) N^* as a function of a_H



(b) N^* as a function of μ_H

Figure 2: Illustration for Proposition 5..4. Parameter values: $A_B = A_S = 0, r = 5\%$, $\gamma = 10\%, 8\%, 6\%, C = 2\% \frac{Nu}{r}, R/u = 50\%, u = 1, (a_L, \mu_L, \sigma_L) = (0.2, 0.04, 0.1), (a_H, \mu_H, \sigma_H) = (0.2, 0.04, 0.1)$ (except for the varying parameter).

References

Allen, F., Gale, D., 1994. Financial Innovation and Risk Sharing. MIT Press, Massachusetts.

Ashcraft, A. B., Schuermann, T., 2009. Understanding the securitization of subprime mortgage credit. Unpublished Working Paper, Federal Reserve Bank of New York.

Axelson, U., 2007. Security design with investor private information. Journal of Finance 62, 2587-2632.

Biais, B., Mariotti, T., 2005. Strategic liquidity supply and security design. The Review of Economic Studies 72, 615-649.

Black, F., Cox, J. C., 1976. Valuing corporate securities: some effects of bond indenture provisions. Journal of Finance 31, 351-367.

Bolton, P., Freixas, X., Shapiro, J., 2012. The credit ratings game. Journal of Finance 67, 85-112.

Boot, A. W. A., Thakor, A. V., 1993. Security design. Journal of Finance 48, 1349-1378.

David, H., Nagaraja H., 2003. Order Statistics. Wiley-Interscience, New Jersey.

DeMarzo, P., Duffie, D., 1999. A liquidity-based model of security design. Econometrica 67, 65-99.

DeMarzo, P., 2005. The pooling and tranching of securities: a model of informed intermediation. Review of Financial Studies 18, 1-35.

DeMarzo, P., Fishman, M. 2007. Optimal long-term financial contracting. Review of Financial Studies 20, 2079-2128.

DeMarzo, P., Sannikov, Y., 2006. Optimal security design and dynamic capital structure in a continuous-time agency model. Journal of Finance 61, 2681-2724.

Diamond, D., Rajan, R., 2000. A theory of bank capital. Journal of Finance 55, 2431-2465.

Downing, C., Jaffee, D., Wallace, N., 2009. Is the market for mortgage-backed securities a market for lemons? Review of Financial Studies 22, 2457-2494.

Duffie, D., Eckner, A., Horel, G., Saita, L., 2009. Frailty correlated default. Journal of Finance 64, 2089-2123.

Duffie, D., Lando, D., 2001. Term structures of credit spreads with incomplete accounting information. Econometrica 69, 633-664.

Duffie, D., Rahi, R., 1995. Financial market innovation and security design: an introduction. Journal of Economic Theory 65, 1-42.

Duffie, D., Singleton, K. J., 2003. Credit Risk: Pricing, Measurement, and Management. Princeton University Press, New Jersey.

Fabozzi, F. J., Davis, H., Choudhry, M., 2006. Introduction to Structured Finance. John Wiley & Sons, New Jersey.

Farhi, E., Tirole, J., 2010. Collective moral hazard, maturity mismatch and systemic bailouts. American Economic Review, forthcoming.

Fender, I., Mitchell, J., 2009. The future of securitisation: how to align incentives? Bank of International Settlements Quarterly Review, September, 27-43.

Fulghieri, P., Dmitry L., 2001. Information production, dilution costs, and optimal security design. Journal of Financial Economics 61, 3-41.

Gorton, G., Pennacchi, G., 1990. Financial intermediaries and liquidity creation. Journal of Finance 45, 49-71.

Guiso, L., Sapienza, P., Zingales, L., 2011. The determinants of attitudes towards strategic default on mortgages. Unpublished Working Paper, European University Institute, Northwestern University and University of Chicago.

Hartman-Glaser, B., Piskorski,T., Tchistyi, A., 2011. Optimal securitization with moral hazard. Journal of Financial Economics, forthcoming.

Innes, R., 1990. Limited liability and incentive contracting with ex-ante action choices. Journal of Economic Theory 52, 45-67.

Kane, E. J., 2009. Incentive roots of the securitization crisis and its early mismanagement. Yale Journal on Regulation 26, 101-112.

Keys, B. J., Mukherjee, T., Seru, A., Vig, V., 2010. Did securitization lead to lax screening? evidence from subprime loans 2001-2006. The Quarterly Journal of Economics 125, 307-362.

Kramkov, D., Schachermayer, W., 1999. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. The Annals of Applied Probability 9, 904-950.

Leland, H. E., 1994. Corporate debt value, bonds covenants and optimal capital structure. Journal of Finance 49, 1213-1252.

Lucas, R. E. and Stokey, N. L., 1989. Recursive Methods in Economic Dynamics. Harvard University Press, Massachusetts.

Manso, G., Strulovici, B., Tchistyi, A., 2010. Performance-sensitive debt. Review of Financial Studies 23, 1819-1854.

Mian, A., Sufi, A., 2009. The consequences of mortgage credit expansion: evidence from the 2007 mortgage default crisis. Quarterly Journal of Economics, forthcoming.

Minton, B. A., Stulz, R., Williamson, R., 2009. How much do banks use credit derivatives to hedge loans? Journal of Financial Services Research 35, 1-31.

Partnoy, F., Skeel, D., 2006. Credit derivatives play a dangerous game. Financial Times, July 16.

Pennacchi, G., 1988. Loan sales and the cost of bank capital. Journal of Finance 43, 375-396.

Sannikov, Y., 2008. A continuous-time version of the principal-agent problem. Review of Economic Studies 75, 957-984.

Stulz, R., 2010. Credit default swaps and the credit crisis. Journal of Economic Perspectives 24, 73-92.

Appendix

A Proof of Theorem 3..1

This section is devoted to a proof of Theorem 3..1. For the reader's convenience, we first provide a brief outline of the main arguments.

- To preserve concavity of the problem, it is useful to make a change of variables from the optimal payments x_k to the utility rate $u_S(x_k)$ that the seller derives from these payments. The reason is that, with this change of variables, the participation constraints and incentive compatibility constraints become linear and the problem can therefore be tackled by standard convex programming techniques. In particular, we start with Lemma A.2 and express the risk sharing rule using the changed variables.
- Lemma A.3 derive an expression for the seller's utility.
- We then use standard duality results (Lemma A.4) to prove existence (and uniqueness) of the optimal contract.
- Finally, in subsection A.1, we derive fully explicit necessary and sufficient conditions for the existence of the optimal contract for the case when the effort is binary.

The following lemma provides an explicit expression for the joint density of $(\tau_1, \cdots, \tau_N)$.

Lemma A.1 *The joint density of* $(\tau_1, \cdots, \tau_k)$, $k \leq N$ *conditional on the effort level* e_j *is given by*

$$f_k^{e_j}(\tau_1, \cdots, \tau_k) \;=\; \mathbf{1}_{\tau_1 < \cdots < \tau_k} \frac{N!}{(N-k)!} p_{e_j}(\tau_1) \cdots p_{e_j}(\tau_k) \left(G_{e_j}(\tau_k)\right)^{N-k}. \tag{24}$$

See David and Nagaraja (2003).

We will also need an auxiliary result to characterize the nature of risk sharing between the intermediary and the investor.

Lemma A.2 *Let $w(x) = (u_S)^{-1}(x)$. Then there exists a unique solution $J(x;d)$ to the equation*

$$u_B'\big(d \;-\; w(J(x;d))\big)\, w'(J(x;d)) \;=\; x\,. \tag{25}$$

Note also that $w'(x) = \frac{1}{u_S'(w(x))}$.

Lemma A.3 *The seller's expected utility from the payment of $x_k(t, \tau_{[1,k]})$, conditional on effort e, is given by*

$$\int_{\mathbb{R}^{k+1}} \psi_k^e(\tau_1, \cdots, \tau_k, t)\, e^{-\gamma t}\; u_S(x_k(t, \tau_{[1,k]}))\, d\tau_1 \cdots d\tau_k\, dt \tag{26}$$

where

$$\psi_k^e(\tau_1, \cdots, \tau_k, t) \;\stackrel{def}{=}\; \frac{N!}{(N-k)!} \mathbf{1}_{\tau_1 \le \cdots \le \tau_k \le t} p_e(\tau_1) \cdots p_e(\tau_k)\, (G_e(t))^{N-k}\,. \tag{27}$$

Proof. The proof is not completely straightforward because we must attend to the fact that the probability density may be defective:

$$\lim_{t\to\infty} G_e(t) \;=\; P_e^\infty \;\ge\; 0.$$

We have

$$P[\tau_{k+1} \in dt | \tau_1, \cdots, \tau_k] \;=\; \frac{f_{k+1}^e(\tau_1, \cdots, \tau_k, t)dt}{f_k(\tau_1, \cdots, \tau_k)} \;=\; (N-k)\frac{p_e(t)\,(G_e(t))^{N-k-1}}{(G_e(\tau_k))^{N-k}} dt$$

and therefore, using the identity

$$\int_t^{\infty} p_e(\tau_{k+1}) \left(G_e(\tau_{k+1})\right)^{N-k-1} d\tau_{k+1} \;=\; \frac{(G_e(t))^{N-k} - (P_e^{\infty})^{N-k}}{N-k}, \tag{28}$$

we get that

$$P[\tau_{k+1} = +\infty | \tau_1, \cdots, \tau_k] \;=\; 1 \;-\; \int_{\tau_k}^{+\infty} (N-k) \frac{p_e(t) \left(G_e(t)\right)^{N-k-1}}{(G_e(\tau_k))^{N-k}} dt \;=\; \frac{(P_e^{\infty})^{N-k}}{(G_e(\tau_k))^{N-k}} \,.$$

Therefore, using (28) once again and changing the order of integration, we get:

$$\begin{aligned}
&\int_{\mathbb{R}_+^{k+1}} f_k^e(\tau_1, \cdots, \tau_k) \, (N-k) \frac{p_e(\tau_{k+1}) \left(G_e(\tau_{k+1})\right)^{N-k-1}}{(G_e(\tau_k))^{N-k}} \\
&\times \int_{\tau_k}^{\tau_{k+1}} e^{-\gamma t} \, u_S(x_k(t, \tau_{[1,k]})) dt \, d\tau_1 \cdots d\tau_{k+1} \\
&= \int_{\mathbb{R}^k} f_k^e(\tau_1, \cdots, \tau_k) \int_{\tau_k}^{+\infty} e^{-\gamma t} \, u_S(x_k(t, \tau_{[1,k]})) \frac{(G_e(t))^{N-k} - (P_e^{\infty})^{N-k}}{(G_e(\tau_k))^{N-k}} dt \, d\tau_1 \cdots d\tau_k
\end{aligned} \tag{29}$$

Therefore, the seller's expected utility from the payment of $x_k(t, \tau_{[1,k]})$, conditional on effort e, is

$$\begin{aligned}
&\int_{\mathbb{R}^k} f_k^e(\tau_1, \cdots, \tau_k) \int_{\tau_k}^{+\infty} e^{-\gamma t} \, u_S(x_k(t, \tau_{[1,k]})) \frac{(G_e(t))^{N-k} - (P_e^{\infty})^{N-k}}{(G_e(\tau_k))^{N-k}} dt \, d\tau_1 \cdots d\tau_k \\
&+ \int_{\mathbb{R}_+^k} f_k^e(\tau_1, \cdots, \tau_k) \, \frac{(P_e^{\infty})^{N-k}}{(G_e(\tau_k))^{N-k}} \int_{\tau_k}^{+\infty} e^{-\gamma t} \, u_S(x_k(t, \tau_{[1,k]})) dt \, d\tau_1 \cdots d\tau_k \\
&= \int_{\mathbb{R}^{k+1}} \psi_k^e(\tau_1, \cdots, \tau_k, t) \, e^{-\gamma t} \; u_S(x_k(t, \tau_{[1,k]})) \, dt \, d\tau_1 \cdots d\tau_k \,,
\end{aligned} \tag{30}$$

where we have used the identity

$$\psi_k^e(\tau_1, \cdots, \tau_k, t) \;=\; f_k^e(\tau_1, \cdots, \tau_k) \, \frac{(G_e(t))^{N-k}}{(G_e(\tau_k))^{N-k}} \,. \tag{31}$$

■

Recall that the utility rate of the seller from the contract $\{x_n\}$ is:

$$v_k(t) = u_S\left(x_k(t)\right) \tag{32}$$

and

$$w(z) = (u_S)^{-1}(z) \tag{33}$$

is the inverse of the seller's utility.

We only study the monopolistic case, the other case is analogous. We will use the standard duality approach. Namely, we can introduce the dual function

$$\begin{aligned} \tilde{u}(\mu_{\rm IC}, \mu_{\rm PC}) \;&=\; \max_{\{x_n\}\geq 0} \Big\{ U_B(\{\delta_n - x_n\}, e_j) \\ &+\; \sum_{i\neq j} \mu_{i,\rm IC}\left(U_S(\{x_n\}, e_j) - U_S(\{x_n\}, e_i)\right) \;+\; \mu_{\rm PC}(U_S(\{x_n\}, e_j) - U_S^0) \Big\} \end{aligned} \tag{34}$$

Then, using the same change of variable (to the utility rate) as above, we can transform the optimization problem over $\{x_n\}$ into a concave one, and it follows that the maximum is attained at

$$x_n(t, \tau_{[1,n]}; \{\mu_{\rm IC}, \mu_{\rm PC}\}) \;=\; \mathbf{1}_{\mathcal{I}_n}\, w\left(J\left(\Psi_n(t, \tau_{[1,n]}; \{\mu_{\rm IC}, \mu_{\rm PC}\})\,;\, \delta_n\right)\right)\,. \tag{35}$$

The following result follows by standard duality arguments (see, e.g., Kramkov and Schachermayer (1999)):

Lemma A.4 *The function $\tilde{u}(\mu_{\rm IC}, \mu_{\rm PC})$ is convex. Suppose that it attains a global minimum over $\mathbb{R}_+^{K+1}$ at a point $\{\mu^*_{\rm IC}, \mu^*_{\rm PC}\}$.*

Then, the corresponding contract $x_n(t, \tau_{[1,n]}; \{\mu^*_{\rm IC}, \mu^*_{\rm PC}\})$ is an optimal contract.

Thus, in order to prove existence of the optimal contract, it suffices to show that the function $\tilde{u}(\mu_{\rm IC}, \mu_{\rm PC})$ does attain a minimum. By standard compactness arguments, it

suffices to show that $\tilde{u}(\mu_{\text{IC}}, \mu_{\text{PC}})$ converges to $+\infty$ as $\{\mu_{\text{IC}}, \mu_{\text{PC}}\} \to \infty$. Indeed, pick a sequence $\{\mu_{\text{IC}}, \mu^l_{\text{PC}}\} \to \infty$ converging to infinity. Then, passing to a subsequence if necessary, we may assume that either (a) there exists a $k > 0$ such that $\mu^l_{k,\text{IC}} \to +\infty$ as $l \to \infty$ or (b) $\mu^l_{\text{PC}} \to \infty$ as $l \to \infty$. By (34), choosing a sub-optimal contract $\{x_n\}$ always gives a lower bound for $\tilde{u}$. Let $\{\bar{x}_n\}$ be a contract that is strictly incentive and participation constraint compatible (the set of such contracts in non-empty by assumption). Then, we have

$$\begin{aligned}
\tilde{u}(\mu^l_{\text{IC}}, \mu^l_{\text{PC}}) \;&\geq\; \Big\{U_B(\{\delta_n - \bar{x}_n\}, e_j) \\
&+\; \sum_{i \neq j} \mu_{i,\text{IC}} \left(U_S(\{\bar{x}_n\}, e_j) - U_S(\{\bar{x}_n\}, e_i)\right) \;-\; \mu_{\text{PC}}(U_S(\{\bar{x}_n\}, e_j) - U^0_S)\Big\} \\
&\geq\; U_B(\{\delta_n - \bar{x}_n\}, e_j) \;+\; \mu^l_{k,\text{IC}}(U_S(\{\bar{x}_n\}, e_j) - U_S(\{\bar{x}_n\}, e_k)) \\
&+\; \mu^l_{\text{PC}}(U_S(\{\bar{x}_n\}, e_j) - U^0_S) \;\to\; +\infty
\end{aligned} \tag{36}$$

as $l \to \infty$, and the claim follows.

A.1 The Binary Effort Case: Precise conditions for Existence of the Optimal Contract

In the general case of multiple effort levels, we prove existence of an optimal contract assuming that the set of strongly incentive- and participation-constraints compatible contracts is non-empty. In this subsection, we consider the case of a binary effort $e \in \{e_H, e_L\}$ and derive fully explicit necessary and sufficient conditions for the existence of the optimal contract. We consider the contract implementing high effort level e_H and we use sub(super)-scripts H and L instead of e_H and e_L unless otherwise stated. We also denote $\mathbf{P}_k = \mathbf{P}_{k,e_L,e_H}$.

Let us first consider the case when the buyer has full bargaining power in designing

the contract. The problem for the buyer can then be written as:[34]

$$\max_{\{v_k\}_{k=0,\cdots N}:\ v_k \geq u_S(0)} \sum_{k=0}^{N-1} \int_{\mathbb{R}_+^{k+1}} \psi_k^H(\tau_1, \cdots, \tau_k, t) e^{-rt} u_B\left(\delta_k - w(v_k(t, \tau_{[1,k]}))\right) dt\, d\tau_1 \cdots d\tau_k$$

under the incentive compatibility (IC) constraint

$$\sum_{k=0}^{N-1} \int_{\mathbb{R}_+^{k+1}} (\psi_k^H - \psi_k^L)(\tau_1, \cdots, \tau_k, t) e^{-\gamma t} v_k(t, \tau_{[1,k]}) dt\, d\tau_1 \cdots d\tau_k \geq C$$

with $C \equiv C_H - C_L$ and under the individual rationality (participation) constraint

$$\sum_{k=0}^{N-1} \int_{\mathbb{R}_+^{k+1}} \psi_k^H(\tau_1, \cdots, \tau_k, t) e^{-\gamma t} v_k(t, \tau_{[1,k]}) dt\, d\tau_1 \cdots d\tau_k \geq U_S^0 + C_H$$

for the seller. Clearly, we can rewrite the IC constraint as

$$\sum_{k=0}^{N-1} \int_{\mathbb{R}_+^{k+1}} \psi_k^H \mathbf{P}_{k,e_i,e_j}(\tau_1, \cdots, \tau_k, t) e^{-\gamma t} v_k(t, \tau_{[1,k]}) dt\, d\tau_1 \cdots d\tau_k \geq C_H$$

Denoting the Lagrange multipliers for the two constraints by μ_1 and μ_2, the first order condition for v_k takes the form

$$-e^{-rt} u_B'\left(\delta_k - w(v_k)\right) w'(v_k) + \mu_1 e^{-\gamma t} \mathbf{P}_k + \mu_2 e^{-\gamma t} = 0$$

when the limited liability constraint (LL), $v_k \geq u_S(0)$, is not binding and $v_k = u_S(0)$

[34] Here, $\{v_k \geq u_S(0)\}$ is the transformed limited liability constraint.

otherwise.[35] Thus,

$$x_k(t, \tau_{[1,k]}, \mu_1, \mu_2) = \max\left\{0,\ \ w\left(J\left(\left(e^{(r-\gamma)t}(\mu_1 \mathbf{P}_k + \mu_2)\right); \delta_k\right)\right)\right\}.$$

Define

$$\begin{aligned} g_k(\mu_1, \mu_2) &= \int\limits_{\mathbb{R}_+^{k+1}} \psi_k^H(\tau_1, \cdots, \tau_k, t)\, P_k\, e^{-\gamma t} u_S(x_k(t, \tau_{[1,k]}, \mu_1, \mu_2)) dt\, d\tau_1 \cdots d\tau_k \\ h_k(\mu_1, \mu_2) &= \int\limits_{\mathbb{R}_+^{k+1}} \psi_k^H(\tau_1, \cdots, \tau_k, t)\, e^{-\gamma t} u_S(x_k(t, \tau_{[1,k]}, \mu_1, \mu_2)) dt\, d\tau_1 \cdots d\tau_k \end{aligned} \tag{37}$$

and

$$\begin{aligned} F_1(\mu_1, \mu_2) &= \sum_{k=0}^{N-1} g_k(\mu_1, \mu_2) \ - \ C \\ F_2(\mu_1, \mu_2) &= \sum_{k=0}^{N-1} h_k(\mu_1, \mu_2) \ - \ C_H \ - \ U_S^0. \end{aligned} \tag{38}$$

Here, as above, $C = C_H - C_L$.

Lemma A.5 *We have*

$$F_1(0, \mu_2) \ = \ -C \ < \ 0$$

for any $\mu_2 \geq 0$ *and*

$$F_2(0,0) = \gamma^{-1}\, u_S(0) - U_S^0 - C_H < 0.$$

Proof. Indeed, $F_1(\mu_1, \mu_2) + C$ is the difference between the seller's utilities from the contractual payments, conditional on high and low effort respectively. But because $\mu_1 = 0$, the payments are deterministic and do not depend on the stochastic process of defaults, this difference is zero.

[35] Standard duality arguments imply that these necessary conditions are also sufficient. See Kramkov and Schachermayer (1999).

The last claim follows because, for $\mu_1 = \mu_2 = 0$, the payments to the intermediary are identically zero. ■

Lemma A.6 *We have* $\lim_{x\to\infty} J(x;d) = u_S(d-\ell_B)$ *and* $\lim_{x\to 0} J(x;d) = u_S(\ell_S)$.

Proof. By assumption, w satisfies $w'(u_S(+\infty)) = \infty$ and $w'(u_S(\ell_S)) = 0$. Therefore,

$$\lim_{J\to u_S(d-\ell_B)} u_B'(d-w(J))w'(J) \;=\; +\infty$$

and

$$\lim_{J\to u_S(\ell_S)} u_B'(d-w(J))w'(J) \;=\; 0.$$

■

The following lemma is a straightforward application of the monotone convergence theorem.

Lemma A.7 *Let* $J_{\max} = \lim_{x\to+\infty} J(x;d)$ $(\leq +\infty)$. *Then,*

$$\lim_{\mu_1\to+\infty} F_2(\mu_1,0) \;=\; J_{\max}\sum_{k=0}^{N-1}\int_{\mathbb{R}_+^{k+1}} \psi_k^H(\tau_1,\cdots,\tau_k,t)e^{-\gamma t}\mathbf{1}_{P_k>0}\,dt\,d\tau_1\cdots d\tau_k - U_S^0 - C_H \tag{39}$$

and

$$\lim_{\mu_2\to+\infty} F_2(\mu_1,\mu_2) \;=\; J_{\max}\,\gamma^{-1} - U_S^0 - C_H\,, \tag{40}$$

and

$$\lim_{\mu_1\to+\infty} F_1(\mu_1,0) \;=\; J_{\max}\sum_{k=0}^{N-1}\int_{\mathbb{R}_+^{k+1}} \psi_k^H(\tau_1,\cdots,\tau_k,t)e^{-\gamma t}P_k\,\mathbf{1}_{P_k>0}\,dt\,d\tau_1\cdots d\tau_k - C \tag{41}$$

Lemma A.8 *The function* $F_1(\mu_1,\mu_2)$ *is monotone increasing in* μ_1. *Thus, if* (41) *is positive then there exists a* $\mu_1^* > 0$ *such that* $F_1(\mu_1^*,0) = 0$. *Otherwise,* $F_1(\mu_1,0) < 0$ *for all* $\mu_1 > 0$.

Lemma A.9 *The function $F_2(\mu_1, \mu_2)$ is monotone increasing in μ_2. If (40) is positive, then there exists a unique solution μ_2^* to $F_2(0, \mu_2) = 0$. This is the optimal contract if $F_1(0, \mu_2^*) \geq 0$.*[36]

Furthermore, the function $F_2(\mu_1, 0)$ is monotone increasing in μ_1; therefore, there exists a $\bar{\mu}_1 > \mu_1^*$ such that $F_2(\bar{\mu}_1, 0) = 0$ if and only if (39) is positive.

For any $\mu_1 \in [0, \bar{\mu}_1]$,[37] *there exists a unique $\mu_2 = \zeta(\mu_1)$ solving $F_2(\mu_1, \mu_2) = 0$.*

Proposition A.10 *The optimal contract when the buyer has full power in designing the contract exists if and only if (40) is positive.*

Proof. Indeed, if (40) is negative, the IR constraint for the seller is violated for any contract and therefore the seller will never participate. Suppose first that μ_1^* exists and is finite. Then, if $F_2(\mu_1^*, 0) > 0$ then the optimal contract corresponds to $(\mu_1^*, 0)$.

If $F_1(0, \mu_2^*) > 0$ then the optimal contract corresponds to $(0, \mu_2^*)$.

Otherwise, we know that $F_2(\mu_1, 0) < 0$ for $\mu_1 \in [0, \bar{\mu}_1]$ and hence, by continuity, there exists a solution $\mu_1 \in [0, \bar{\mu}_1]$ to

$$F_1(\mu_1, \zeta(\mu_1)) \;=\; 0\,. \tag{42}$$

because $F_1(0, \zeta(0)) \;=\; F_1(0, \mu_2^*) \;<\; 0$, whereas $F_1(\bar{\mu}_1, \zeta(\bar{\mu}_1)) \;=\; F_1(\bar{\mu}_1, 0) \;\geq F_1(\mu_1^*, 0) =$ 0 and the existence follows immediately from continuity. To prove uniqueness, we show that $F_1(\mu_1, \zeta(\mu_1))$ is monotone increasing in μ_1. Indeed,

$$\frac{\partial \zeta}{\partial \mu_1} \;=\; -\frac{\frac{\partial F_2}{\partial \mu_1}}{\frac{\partial F_2}{\partial \mu_2}}\,. \tag{43}$$

[36] This can never happen if the optimal contract implements the highest effort level, but it can happen if the contract implements a lower effort level.

[37] If $\bar{\mu}_1$ does not exist, we set $\bar{\mu}_1 = \infty$.

Furthermore,

$$\frac{\partial g_k}{\partial \mu_1} \;=\; \int\limits_{\mathbb{R}_+^k} \psi_k^H(\tau_1,\cdots,\tau_k,t) \int_{\tau_k}^{\bar{T}} e^{(r-2\gamma)\,t}\mathbf{P}_k^2\, J_x(e^{(r-\gamma)t}(\mu_1\mathbf{P}_k+\mu_2);d_k)dt\,d\tau_1\cdots d\tau_k \quad (44)$$

and

$$\frac{\partial g_k}{\partial \mu_2} \;=\; \int\limits_{\mathbb{R}_+^k} \psi_k^H(\tau_1,\cdots,\tau_k,t) \int_{\tau_k}^{\bar{T}} e^{(r-2\gamma)\,t}\mathbf{P}_k\, J_x(e^{(r-\gamma)t}(\mu_1\mathbf{P}_k+\mu_2);d_k)dt\,d\tau_1\cdots d\tau_k \quad (45)$$

Similarly,

$$\frac{\partial h_k}{\partial \mu_1} \;=\; \int\limits_{\mathbb{R}_+^k} \psi_k^H(\tau_1,\cdots,\tau_k,t) \int_{\tau_k}^{\bar{T}} e^{(r-2\gamma)\,t}\mathbf{P}_k\, J_x(e^{(r-\gamma)t}(\mu_1\mathbf{P}_k+\mu_2);d_k)dt\,d\tau_1\cdots d\tau_k \;=\; \frac{\partial g_k}{\partial \mu_2} \quad (46)$$

and

$$\frac{\partial h_k}{\partial \mu_2} \;=\; \int\limits_{\mathbb{R}_+^k} \psi_k^H(\tau_1,\cdots,\tau_k,t) \int_{\tau_k}^{\bar{T}} e^{(r-2\gamma)\,t}\, J_x(e^{(r-\gamma)t}(\mu_1\mathbf{P}_k+\mu_2);d_k)dt\,d\tau_1\cdots d\tau_k \quad (47)$$

Therefore,

$$\frac{d}{d\mu_1}F_1(\mu_1,\zeta(\mu_1)) \;=\; \frac{\partial F_1}{\partial \mu_1} \;-\; \frac{\partial F_1}{\partial \mu_2}\frac{\frac{\partial F_2}{\partial \mu_1}}{\frac{\partial F_2}{\partial \mu_2}} \quad (48)$$

and hence, to prove monotonicity, we need to show that

$$\frac{\partial F_2}{\partial \mu_2}\frac{\partial F_1}{\partial \mu_1} \;\geq\; \frac{\partial F_2}{\partial \mu_1}\frac{\partial F_1}{\partial \mu_2}\,. \quad (49)$$

We use the following slight modification of the Cauchy-Schwarz inequality.

Lemma A.11 *For any function $\psi_k(x), g_k(x)$, we have*

$$\sum_k \int_{\mathbb{R}^m} \psi_k^2(x)dx \sum_k \int_{\mathbb{R}^m} g_k^2(x)dx \;\geq\; \left(\sum_k \int_{\mathbb{R}^m} \psi_k(x)g_k(x)dx\right)^2.$$

The required inequality (49) follows now from (44)-(46) and Lemma A.11 if we let

$$\begin{aligned} \psi_k^2(x,\tau_k) &= \psi_k^H(\tau_1,\cdots,\tau_k,t)\,\mathbf{1}_{[\tau_k,\bar{T}]}\,e^{(r-2\gamma)\,t}\mathbf{P}_k^2 J_x(e^{(r-\gamma)t}(\mu_1\mathbf{P}_k+\mu_2);d_k) \\ g_k^2(x,\tau_k) &= \psi_k^H(\tau_1,\cdots,\tau_k,t)\,\mathbf{1}_{[\tau_k,\bar{T}]}\,e^{(r-2\gamma)\,t} J_x(e^{(r-\gamma)t}(\mu_1\mathbf{P}_k+\mu_2);d_k) \end{aligned} \tag{50}$$

Finally, the fact that the maturity of the contract is finite follows directly from $\gamma > r$.

■

Now, let us consider the case where the seller has all the bargaining power. For simplicity, we assume that $\ell_B = -\infty$.

Then, F_1 stays the same, but F_2 is replaced by

$$\tilde{F}_2(\mu_1,\mu_2) \;=\; \sum_{k=0}^{N-1} \tilde{h}_k(\mu_1,\mu_2) \;-\; U_B^0,$$

where

$$\tilde{h}_k(\mu_1,\mu_2) \;=\; \int\limits_{\mathbb{R}_+^{k+1}} \psi_k^H(\tau_1,\cdots,\tau_k,t)e^{-rt}u_B\left(d_k - w(v_k(t,\tau_{[1,k]}))\right)dt\,d\tau_1\cdots d\tau_k$$

The same arguments already given imply that the following is true.

Lemma A.12 *The function $\tilde{F}_2$ is monotone decreasing in μ_2 and $\tilde{F}_2 \to -\infty$ when $\mu_2 \to +\infty$ and there exists a unique solution $\tilde{\mu}_2^*$ to $\tilde{F}_2(0,\tilde{\mu}_2^*) = 0$. This is the optimal contract if $F_1(0,\tilde{\mu}_2^*) \geq 0$.*[38]

Furthermore, $\tilde{F}_2(\mu_1,0)$ is monotone decreasing in μ_1 and therefore there exists a

[38]Note that this can only happen if we are not implementing the highest effort level.

$\bar{\mu}_1 > 0$ *such that* $\tilde{F}_2(\bar{\mu}_1, 0) = 0$. *For any* $\mu_1 \in [0, \bar{\mu}_1]$, $\tilde{F}_2(\mu_1, 0) > 0$ *and, there exists, therefore, a unique* $\mu_2 = \tilde{\zeta}(\mu_1)$ *such that* $\tilde{F}_2(\mu_1, \tilde{\zeta}(\mu_1)) = 0$.

Note that, by definition, $\tilde{\zeta}(\bar{\mu}_1) = 0$ and $\tilde{\zeta}(0) = \mu_2^*$.

Proposition A.13 *For the case when the seller has full bargaining power in designing the contract, the optimal contract exists if and only if at least one of the two numbers* $F_1(0, \mu_2^*) < 0$ *and* $F_1(\bar{\mu}_1, 0) < 0$ *is positive.*

Proof. If $F_1(0, \mu_2^*) > 0$ then $(0, \mu_2^*)$ corresponds to the optimal contract.

Suppose now that $F_1(0, \mu_2^*) < 0$, that is $F_1(0, \tilde{\zeta}(0)) < 0$. Using (25), it is not difficult to show that $F_1(\mu_1, \tilde{\zeta}(\mu_1))$ is monotone increasing in μ_1. Now, the claim follows if $F_1(\bar{\mu}_1, 0) > 0$.

However, if $F_1(\bar{\mu}_1, 0) < 0$ and $F_1(0, \mu_2^*) < 0$, then the optimal contract obviously does not exist. Indeed, the unique pair (μ_1, μ_2) has $\mu_2 = 0$ in this case, but this contract does not satisfy the optimality conditions for any finite $\tilde{\mu}_{\text{PC}}$. ■

B Hazard Rate and Likelihood Ratio Properties

Proof of Proposition 3..3. The claim of Proposition 3..3 directly follows from the following lemma.

Lemma B.1 *Suppose that the densities* $p_{e_i} \prec_{hr} p_{e_j}$. *Then, for any* $k \geq 0$ *and any* $\tau_1 \leq \cdots \leq \tau_k \leq \tau_{k+1} \leq t$,

$$\mathbf{P}_{k,e_i,e_j}(t; \tau_{[1,k]}) \;\geq\; \mathbf{P}_{k+1}(t; \tau_{[1,k+1]})\,.$$

Proof of Lemma B.1. By definition, we need to show that

$$\frac{p_{e_i}(\tau_1)\cdots p_{e_i}(\tau_k)(G_{e_i}(t))^{N-k}}{p_{e_j}(\tau_1)\cdots p_{e_j}(\tau_k)(G_{e_j}(t))^{N-k}} \;\leq\; \frac{p_{e_i}(\tau_1)\cdots p_{e_i}(\tau_k)p_{e_i}(\tau_{k+1})(G_{e_i}(t))^{N-k-1}}{p_{e_j}(\tau_1)\cdots p_{e_j}(\tau_k)p_{e_j}(\tau_{k+1})(G_{e_j}(t))^{N-k-1}}\,,$$

that is

$$\frac{G_{e_i}(t)}{G_{e_j}(t)} \;\leq\; \frac{p_{e_i}(\tau_{k+1})}{p_{e_j}(\tau_{k+1})}\,.$$

Because, $p_{e_i} \prec_{hr} p_{e_j}$, we have

$$\frac{p_{e_i}(\tau_{k+1})}{p_{e_j}(\tau_{k+1})} \;>\; \frac{G_{e_i}(\tau_{k+1})}{G_{e_j}(\tau_{k+1})}\,.$$

Thus, it remains to be shown that $G_{e_i}(t)/G_{e_j}(t)$ is monotone decreasing in t. This follows immediately from $p_{e_i} \prec_{hr} p_{e_j}$ and from the identity

$$\frac{d}{dt}\left(\frac{G_{e_i}(t)}{G_{e_j}(t)}\right) \;=\; \frac{-p_{e_i}(t)G_{e_j}(t) \;+\; p_{e_j}(t)G_{e_i}(t)}{(G_{e_j}(t))^2}\,.$$

■

■

Proof of Proposition 6..1. The claims about $\prec_{lr}$ order follow by direct calculation. A direct calculation also shows that $p^{a,m,\sigma}$ is monotone increasing in x in the sense of $\prec_{lr}$, and that $P^{\infty}_{a,m,\sigma}$ is increasing in x, m and decreasing in σ. So, the claim for x follows. Indeed, let $x_1 \geq x_2$ and let $p_i(t) = p^{x_i,m,\sigma}(t)\ ,\ i = 1, 2$, and $G_i(t) = G^{x_i,m,\sigma}(t),\ i = 1, 2$ and $P^{\infty}_i = \lim_{t\to\infty} G_i(t)$. Then, by the monotone likelihood property of the densities, we have

$$\int_t^{\infty} p_1(s)ds \;=\; \int_t^{\infty} \frac{p_1(s)}{p_2(s)}\, p_2(s)ds \;\geq\; \frac{p_1(t)}{p_2(t)} \int_t^{\infty} p_2(s)ds\,.$$

Therefore,

$$\frac{p_1(t)}{G_1(t)} \;=\; \frac{p_1(t)}{\int_t^{\infty} p_1(s)ds \;+\; P^{\infty}_1} \;\leq\; \frac{p_2(t)}{\int_t^{\infty} p_2(s)ds \;+\; \frac{p_2(t)\,P^{\infty}_1}{p_1(t)}}\,.$$

Thus, to complete the proof, it remains to be shown that

$$\frac{p_2(t)\,P_1^\infty}{p_1(t)} \;\geq\; P_2^\infty\,. \tag{51}$$

Because $p_2(t)/p_1(t)$ is decreasing, we have

$$\frac{p_2(t)\,P_1^\infty}{p_1(t)} \;\geq\; \lim_{t\to\infty}\frac{p_2(t)\,P_1^\infty}{p_1(t)} \;=\; \frac{x_2}{x_1}\,e^{\frac{m(x_1-x_2)}{\sigma^2}}(1-e^{-\frac{2mx_1}{\sigma^2}})$$

and a direct calculation implies that the required inequality (51) is equivalent to the monotonicity of the function $x^{-1}\,(e^x-e^{-x})$ for $x\geq 0$. This follows because

$$x^{-1}\,(e^x-e^{-x}) \;=\; \sum_{n=0}^{\infty}\frac{x^{2n}}{(2n+1)!}\,.$$

Now, for the m parameter, let $p_{e_i}(t)=p^{a,m_{e_i},\sigma}$, $p_{e_j}(t) \;=\; p^{a,m_{e_j},\sigma}$. Then, a direct calculation shows that $p_{e_j}\prec_{lr} p_{e_i}$. It also follows from the definition of the stopping time that $G_{e_j}(t)\geq G_{e_i}(t)$ for all t. Indeed, $G_{e_j}(t)$ is the probability that a geometric Brownian motion A_t^j that starts at A_0 does not fall below the barrier $A_B<A_0$ over the time interval $[0,t]$. But since $m_j>m_i$, we have

$$A_t^j \;=\; A_0\,e^{m_jt+\sigma B_t} \;\geq\; A_0\,e^{m_it+\sigma B_t} \;=\; A_t^i,$$

and therefore A_t^j will always fall below A_B later than A_t^i.

Now, let

$$\phi(t) \;=\; \frac{p_{e_i}(t)}{p_{e_j}(t)} \;-\; \frac{G_{e_i}(t)}{G_{e_j}(t)}\,.$$

Then, our goal is to show that $\phi(t)\geq 0$ for all $t\geq 0$. We have $\phi(0) \;=\; 0$ and, since $\frac{p_{e_i}(t)}{p_{e_j}(t)}$

is increasing, we always have $p_{e_i}(t) > p_{e_j}(t)$. Furthermore, a direct calculation shows that

$$-\frac{d}{dt}\left(\frac{G_{e_i}(t)}{G_{e_j}(t)}\right) \;=\; \frac{p_{e_i}(t)G_{e_j}(t) - p_{e_j}(t)G_{e_i}(t)}{G^2_{e_j}(t)}$$

and therefore,

$$p_{e_i}(t)G_{e_j}(t) - p_{e_j}(t)G_{e_i}(t) \;\geq\; G_{e_i}(t)\left(p_{e_i}(t) - p_{e_j}(t)\right) \;\geq\; 0\,.$$

Finally, the claim for the volatility follows from scale invariance: If $\sigma_{e_j} = \alpha\sigma_{e_i}$ for some $\alpha > 1$, we have

$$p^{x,\sigma^H,m} \;=\; p^{x/\alpha,\sigma^L,m/\alpha}$$

, and the claim follows from the just proven results for the (x, m) parameters. ■

The following result is a direct consequence of the above.

Proposition B.2 *Suppose that $p_{e_j} = p^{(a_j,m_j,\sigma_j)}$: that is, exerting effort level e_j leads to selecting borrowers with parameters (a_j, m_j, σ_j). Then, higher effort reduces default risk if and only if $\frac{a_j}{\sigma_j}$ and $\frac{m_j a_j}{\sigma_j^2}$ is monotone increasing in j. In this case, for any $j > i$, we have:*

(1) If $\frac{m_j}{\sigma_j} > \frac{m_i}{\sigma_i}$, then $p_{e_i} \prec_{hr} p_{e_j}$, and there exists a $\bar{t} > 0$ such that the likelihood ratio $p_{e_j}(t)/p_{e_i}(t)$ is monotone increasing for $t < \bar{t}$ and monotone decreasing for $t > \bar{t}$.[39] Thus, higher effort reduces the default hazard rate, but its effect on the likelihood ratio is ambiguous.

(2) If $\frac{m_j}{\sigma_j} < \frac{m_i}{\sigma_i}$, then $p_{e_i} \prec_{lr} p_{e_j}$, and there exists a threshold $\hat{t} > 0$ such that the hazard rates satisfy $h_{e_i}(t) \;\geq\; h_{e_j}(t)$ if and only if $t < \hat{t}$. Thus, higher effort reduces the hazard rate for $t < \hat{t}$, but increases it for $t > \hat{t}$.

[39] In fact, $\bar{t} = \frac{(a_j/\sigma_j)^2 - (a_i/\sigma_i)^2}{(m_j/\sigma_j)^2 - (m_i/\sigma_i)^2}$.

Proposition B.3 *Suppose that, as in Proposition B.2, $p_{e_k} = p^{(a_k,m_k,\sigma_k)}$ for some $a_k, m_k, \sigma_k > 0$ and that higher effort reduces default risk. Then, the following is true:*

(1) If $\frac{m_j}{\sigma_j} \geq \frac{m_i}{\sigma_i}$ for all i, then the optimal contract payments $x_n(t, \tau_{[1,n]})$ are

- *decreasing with n; and*
- *increasing (decreasing) in $\tau_k, k = 1, \cdots, n$ for sufficiently small (large) τ_k.*

(2) If $\frac{m_j}{\sigma_j} \leq \frac{m_i}{\sigma_i}$ for all i then the optimal contract payments $x_n(t, \tau_{[1,n]})$ are

- *decreasing (increasing) in n for sufficiently small (large) n; and*
- *increasing in $\tau_k, k = 1, \cdots, n$.*

Proof of Proposition 4..4. The proof follows directly both from Theorem 4..3 and from the following auxiliary result:

Lemma B.4 *For any real numbers $a_1, \cdots, a_K, \alpha_1, \cdots, \alpha_K \in \mathbb{R} \setminus \{0\}$, equation*

$$\sum_i a_i \, e^{\alpha_i t} \;=\; 0 \tag{52}$$

can have at most $K-1$ solutions whereas equation

$$\sum_i a_i \, e^{\alpha_i t} \;=\; A \tag{53}$$

can have at most K solutions for any $A \neq 0$.

Proof of Lemma B.4. The proof is by induction. For $K = 1$, the claim is obvious. Suppose now that the claim is proven for $K = M$, and let us show it for $K = M + 1$. Suppose, on the contrary, that (52) has at least $M + 1$ solutions. Then, the same is true for

$$\sum_{i=2}^{M+1} a_i \, e^{(\alpha_i - \alpha_1)\, t} \;=\; -a_1, \tag{54}$$

which is impossible because, by the induction hypothesis, it may have at most M solutions. (53) follows directly from Roll Theorem and the above results. ■ ■

C Proof of Theorems 4..1 and 4..3

For simplicity, we only consider the case when the buyer is risk neutral, $A_B = 0$, and we assume that the effort is binary, $e \in \{e_H, e_L\}$, and that the investor has all the bargaining power.

We start with the following important observation.

Lemma C.1 *Suppose that the investor is risk neutral (i.e., $A_B = 0$). If the desired effort level is not minimal, then at least one of the IC constraints is binding.*

Proof. Indeed, if none of the constraints are binding, we get that the contract payments rates $\{x_n\}$ are deterministic and therefore it is always optimal for the issuer to exert lowest possible effort. ■

In this case, a direct calculation shows that

$$w(z) = (u_S)^{-1} = -A_S^{-1} \ln(1 - A_S z)$$

and

$$J(z; d) = A_S^{-1} \left(1 - z^{-1}\right)$$

is independent of d. Hence, the LL constraint is not binding if and only if

$$e^{(r-\gamma)t}(\mu_1 \mathbf{P}_k + \mu_2) \ \geq \ 1 \, .$$

We also assume that both IC and IR constraints are binding.[40] Denote

$$\mathbf{1}_k \;\equiv\; \mathbf{1}_{e^{(r-\gamma)t}(\mu_1 \mathbf{P}_k + \mu_2) \;\geq\; 1}\,.$$

Then, we are considering the system of equations

$$\begin{aligned} \phi_1(\mu_1,\mu_2) \;&=\; A_S\, C \\ \phi_2(\mu_1,\mu_2) \;&=\; A_S(U_S^0 + C) \end{aligned} \tag{55}$$

with

$$\begin{aligned} &\phi_1(\mu_1,\mu_2) \\ &= \sum_{k=0}^{N-1} \int_{\mathbb{R}_+^k} \int_{\tau_k}^{\bar{T}} \psi_k^H(\tau_1,\cdots,\tau_k,t)\, \mathbf{P}_{k,e_i,e_j}\, e^{-\gamma\, t} \mathbf{1}_k \left(1 - \frac{1}{e^{(r-\gamma)t}(\mu_1 \mathbf{P}_k + \mu_2)}\right) dt\, d\tau_1 \cdots d\tau_k \\ &\phi_2(\mu_1,\mu_2) = \sum_{k=0}^{N-1} \int_{\mathbb{R}_+^k} \int_{\tau_k}^{\bar{T}} \psi_k^H(\tau_1,\cdots,\tau_k,t)\, e^{-\gamma\, t} \mathbf{1}_k \left(1 - \frac{1}{e^{(r-\gamma)t}(\mu_1 \mathbf{P}_k + \mu_2)}\right) dt\, d\tau_1 \cdots d\tau_k \end{aligned} \tag{56}$$

Interestingly enough, the only dependence on A_S is through the right-hand sides of (55). Let us determine the dependence of μ_1, μ_2 on A_S to calculate the limit as $A_S \to 0$. We have

$$\phi_2(\mu_1, \zeta(\mu_1, A_S)) \;=\; A_S(U_S^0 + C)\,,$$

which gives

$$\frac{\partial \zeta}{\partial A_S} \;=\; \frac{U_S^0 + C}{\frac{\partial \phi_2}{\partial \mu_2}}.$$

Thus, differentiating

$$\phi_1(\mu_1, \zeta(\mu_1, A_S)) \;=\; A_S\, C$$

[40] This is indeed always true in the risk neutral limit. The general case follows by a modification of the arguments below.

we get

$$\frac{\partial \mu_1}{\partial A_S} \;=\; \frac{C - \frac{\partial \phi_1}{\partial \mu_2}\frac{\partial \zeta}{\partial A_S}}{\frac{\partial \phi_1}{\partial \mu_1} + \frac{\partial \phi_1}{\partial \mu_2}\frac{\partial \zeta}{\partial \mu_1}} \;=\; \frac{\left(\frac{\partial \phi_2}{\partial \mu_2} - \frac{\partial \phi_1}{\partial \mu_2}\right) C - U_S^0 \frac{\partial \phi_1}{\partial \mu_2}}{\frac{\partial \phi_1}{\partial \mu_1}\frac{\partial \phi_2}{\partial \mu_2} - \frac{\partial \phi_1}{\partial \mu_2}\frac{\partial \phi_2}{\partial \mu_1}}$$

Lemma C.2 *Both μ_1, μ_2 stay bounded when $A_S \to 0$.*

Proof. Because $\phi_2(\mu_1, 0)$ is monotone increasing in μ_1, we immediately get that the solution $\bar{\mu}_1$ to $\phi_2(\bar{\mu}_1, 0) = A_S(U_S^0 + C)$ is monotone increasing in A_S. The same is true for μ_1^* and μ_2^*. Thus, if only one of the IC or IR constraints is binding, we are done.

Suppose now that both constraints are binding. Because $\mu_1 \in [0, \bar{\mu}_1]$, then μ_1 stays bounded as $A_S \to 0$. Furthermore, it is straightforward to show that $\phi_2(\mu_1, \mu_2)$ converges to a positive number when μ_1 stays bounded and $\mu_2 \to \infty$, and it immediately follows that μ_2 has to stay bounded when $A_S \to 0$. ■

Lemma C.3 *For sufficiently small A_S, we have $x_k = 0$ for $k \geq 1$.*

Proof. By Proposition 3..3, x_k is monotone decreasing with k and therefore $\max x_0(t) > 0$ because otherwise the whole payment stream will be identically zero.

Let us now show that $\max_t A_S\, x_0(t) \to 0$ as $A_S \to 0$. For simplicity, we only consider the case where the IR constraint is binding. The case where only IC is binding is completely analogous.

Using the inequality

$$A_S^{-1}(1 - e^{-A_S x}) \;>\; x$$

for all $x \geq 0$, we get that

$$\begin{aligned} U_S^0 \;&\geq\; \sum_{k=0}^{N-1} \int_{\mathbb{R}_+^k} \int_{\tau_k}^{\bar{T}} \psi_k^H(\tau_1, \cdots, \tau_k, t)\, e^{-\gamma t} x_k(t, \tau_{[1,k]}) dt\, d\tau_1 \cdots d\tau_k \\ &\geq\; \int_{\tau_k}^{\bar{T}} \psi_0^H(t)\, e^{-\gamma t} x_0(t)\, dt \end{aligned} \tag{57}$$

where

$$x_0 \;=\; \max\{0, \frac{1}{A_S}\log\left(e^{(r-\gamma)t}(\mu_1\mathbf{P}_0+\mu_2)\right)\}\,.$$

Now, suppose that $\max_t A_S\, x_0(t) \not\to 0$. Then, for some $\varepsilon > 0$, there exists a sequence $A_{Sn} \to 0$ such that the corresponding contractual payments satisfy $\max A_{Sn} x_0(t,n) > \varepsilon$. Because, by Lemma C.2 μ_1, μ_2 stay uniformly bounded, the threshold

$$\bar{T}_0 \;\equiv\; \sup\{t>0:\; x_0(t)>0\}$$

also stays uniformly bounded. Therefore, the maximum location t_0 of x_0 also stays uniformly bounded, and so does the derivative of $A_S x_0$ in $(\max\{0,t_1(0)\},\bar{T})$. Let $K = \sup_{(\max\{0,t_1(0)\},\bar{T})} A_S|x_0'(t)|$. Then, for all $t \in (t_0, t_0+\varepsilon/(2K))$ we have

$$A_S x_0(t) \;=\; A_S x_0(t_0) + \int_{t_0}^{t} A_S x_0'(s)ds \;\geq\; \varepsilon/2\,.$$

Therefore,

$$U_S^0 \;\geq\; 0.5\varepsilon\, A_S^{-1}\int_{t_0}^{t_0+\varepsilon/(2K)} \psi_0^H(t)e^{-\gamma t}dt,$$

which is impossible because the right-hand side converges to $+\infty$ when $A_S \to 0$. Thus,

$$\max\left\{e^{(r-\gamma)t}\left(\mu_{\text{PC}} \;+\; \sum_{i\neq j}\mu_{\text{IC},i}\mathbf{P}_{0,e_i,e_j}(t)\right)\right\} \;\to\; 1 \tag{58}$$

when $A_S \to 0$.

Because $x_k \leq x_1 \;\leq\; x_0$, it suffices to show that x_1 is identically zero for small A_S. Because $\bar{T}_0$ stays bounded when $A_S \to 0$, the support of x_1 is uniformly bounded from above by some $\tilde{T}$ as $A_S \to 0$. Furthermore, by the argument from the proof of Lemma B.1,

$$\mathbf{P}_{1,e_i,e_j} \;\leq\; 1 \;-\; \left(\frac{G_{e_i}(t)}{G_{e_j}(t)}\right)^{N-1}\frac{p_{e_i}(t)}{p_{e_j}(t)}$$

By continuity and the fact that e_j has the highest hazard rate, we have

$$\min_{i \neq j} \inf_{(0,\tilde{T}]} \frac{p_{e_i}(t) G_{e_j}(t)}{p_{e_j}(t) G_{e_j}(t)} \;=\; \kappa \;>\; 1$$

Furthermore,

$$\min_{i} \min_{t \in [0,\tilde{T}]} \mathbf{P}_{1,e_i,e_j} \;\equiv\; \delta > 0.$$

We consider two cases.

Case 1. In the limit $A_S \to 0$, $\min_i \mu_{\mathrm{IC},i} > \varepsilon$ for some $\varepsilon > 0$. In this case, because

$$\mu_{\mathrm{PC}} \;+\; \sum_{i \neq j} \mu_{\mathrm{IC},i} \;\leq\; M$$

, for some $M > 0$, we get that

$$e^{(r-\gamma)t} \left(\mu_{\mathrm{PC}} \;+\; \sum_{i \neq j} \mu_{\mathrm{IC},i} \mathbf{P}_{0,e_i,e_j}(t) \right) \;\leq\; \alpha\, e^{(r-\gamma)t} \left(\mu_{\mathrm{PC}} \;+\; \sum_{i \neq j} \mu_{\mathrm{IC},i} \mathbf{P}_{0,e_i,e_j}(t) \right)$$

where $\alpha < 1$ but sufficiently close to 1 so that

$$\alpha \;+\; (\varepsilon\delta)^{-1}(1-\alpha)\, M \;<\; \kappa.$$

Thus, by (58),

$$e^{(r-\gamma)t} \left(\mu_{\mathrm{PC}} \;+\; \sum_{i \neq j} \mu_{\mathrm{IC},i} \mathbf{P}_{0,e_i,e_j}(t) \right) \;<\; 1$$

for sufficiently small A_S, and the claim follows.

Case 2. Suppose that there is a sequence $A_{Sn} \to 0$ such that the corresponding $\mu_{\mathrm{IC},i}(n) \to 0$ for all i. By (58), this immediately yields $\mu_{PC} \to 1$. Therefore, arguments of all local maxima (if there are any) of $x_0(t)$ will converge to 0 when $n \to \infty$. Thus, all payments will be concentrated at time zero in the limit; thus, in the limit, the utility

of the intermediary will be independent of the effort level. This result would mean that the IC constraints cannot be binding, which is a contradiction. ■

To complete the proof, we note that, since Lagrange multipliers stay bounded, we can always pick a convergent subsequence of those. For this subsequence, it follows from the above arguments that the contract converges to a sum of delta functions at the locations of those local maxima of x_0 for which

$$e^{(r-\gamma)t}\left(\mu_{\mathrm{PC}} \; + \; \sum_{i\neq j}\mu_{\mathrm{IC},i}\mathbf{P}_{0,e_i,e_j}(t)\right)$$

converges to 1.

Finally, in the binary effort case, we note that it is possible to show that Lagrange multipliers are monotone decreasing in A_S and therefore converge to finite limits.

D Proof of Theorem 4..5

Here, we prove the following more general statement.

Theorem D.1 *Suppose that the effort is binary, default time distributions are from the Black and Cox model, $p_{e_L} \prec_{hr} p_{e_H}$ and the desired effort level is e_H. Then, in the risk neutral limit, the optimal contract makes a lump sum payment $y_0 \geq 0$ to the intermediary at time 0, and then a lump sum payment $y_1 > 0$ at a time t^* if no default occurs before $t = t^*$. Furthermore:*

(1) Suppose that the investor has full bargaining power in designing the contract (monopolistic case). Then, there exist thresholds $C^*_{B,1} < C^*_{B,2}$ such that:

(i) If $C_H < C^*_{B,1}$, we have $y_0 > 0$ and $t^* = t^*_1$ is independent of C_H.

(ii) If $C^*_{B,1} \leq C_H < C^*_{B,2}$, we have $y_0 = 0$ and t^* is monotone increasing in C_H.

(iii) If $C_H \geq C^*_{B,2}$, we have $y_0 = 0$ and $t^* = t^*_0$ is independent of C_H.

(2) Suppose that the intermediary has full bargaining power in designing the contract (competitive case). Then, there exists a threshold C_S^* such that:

(i) If $C_H < C_S^*$, we have $y_0 > 0$ and $t^* = t_1^*$ is independent of C_H.

(ii) If $C_H \geq C_S^*$, we have $y_0 = 0$ and t^* is monotone increasing in C_H.

Proof. By Theorem 4..3, we know that the optimal contract makes at most two payments, at time zero and at a time $t^* > 0$. Thus, determining the optimal contract is equivalent to finding the contract in this class, maximizing the utility of the agent who has full bargaining power.

Because the desired effort level is e_H, the IC constraint is binding and therefore we always have

$$y_1 = Y(t^*),$$

where

$$Y(t) \;=\; \frac{e^{\gamma t}\,(C_H - C_L)}{(G_{e_H}(t))^N - (G_{e_L}(t))^N}\,. \tag{59}$$

Therefore:

$$\begin{aligned} U_S(\{dX\}, e_i) \;&=\; y_0 + \text{Prob}[\tau_1 > t^* | e_i]\, e^{-\gamma t^*}\, y_1 \;-\; C_i \\ &=\; y_0 + (G_{e_i}(t^*))^N\, e^{-\gamma t^*}\, y_1 \;-\; C_i \end{aligned} \tag{60}$$

(1) Suppose that the investor has full bargaining power. Then, the investor's maximization problem takes the form

$$\min_{t, y_0} \left\{ y_0 + \frac{e^{(\gamma - r)t}\,(C_H - C_L)}{1 - (G_{e_L}(t)/G_{e_H}(t))^N} \;:\; y_0 \geq 0,\; y_0 + \frac{(C_H - C_L)}{1 - (G_{e_L}(t)/G_{e_H}(t))^N} \geq U_S^0 + C_H \right\}$$

Because $p_{e_L} \prec_{hr} p_{e_H}$, the quotient $\frac{G_{e_H}(t)}{G_{e_L}(t)}$ is monotone increasing in t; therefore, a direct calculation shows that

$$U_S^0 + C_H \;\leq\; \frac{C_H - C_L}{1 - (G_{e_L}(t^*)/G_{e_H}(t^*))^N} \tag{61}$$

is equivalent to

$$t^* \leq t_B^* \,,$$

where $t_B^* \;=\; t_B^*$ is the unique solution to

$$\frac{G_{e_H}(t_B^*)}{G_{e_L}(t_B^*)} \;=\; \left(\frac{U_S^0 + C_H}{U_S^0 + C_L}\right)^{1/N} \,. \tag{62}$$

Note that this unique solution only exists if

$$\frac{P_{e_H}^\infty}{P_{e_L}^\infty} \;>\; \left(\frac{U_S^0 + C_H}{U_S^0 + C_L}\right)^{1/N} \,.$$

Otherwise, we set $t_B^* = +\infty$. Clearly, $t_B^* \to 0$ as $N \to \infty$.

Thus, if $t < t_B^*$, the optimal choice is clearly $y_0 = 0$ and the cost minimization problem takes the form

$$\min_{t \leq t_B^*} \frac{e^{(\gamma - r)t}\,(C_H - C_L)}{1 - (G_{e_L}(t)/G_{e_H}(t))^N} \,, \tag{63}$$

and the minimum is clearly attained at $\min\{t_B^*, t_0^*\}$.

If $t > t_B^*$, y_0 needs to be positive to satisfy the IR constraint of the seller, and the optimal choice is clearly

$$y_0 \;=\; U_S^0 + C_H - \frac{(C_H - C_L)}{1 - (G_{e_L}(t)/G_{e_H}(t))^N} \,.$$

Therefore, the cost minimization problem takes the form

$$\min_{t \geq t_B^*} \left(U_S^0 + C_H + \frac{(C_H - C_L)(e^{(\gamma - r)t} - 1)}{1 - (G_{e_L}(t)/G_{e_H}(t))^N} \right) \,, \tag{64}$$

and the minimum is clearly attained at $\max\{t_B^*, t_1^*\}$. The minimal cost is then given by the minimum of the two quantities (70) and (71).

Furthermore, a direct calculation shows that

$$\frac{d}{dt}\frac{(C_H - C_L)(e^{(\gamma-r)t}-1)}{1-(G_{e_L}(t)/G_{e_H}(t))^N} > \frac{d}{dt}\frac{e^{(\gamma-r)t}(C_H - C_L)}{1-(G_{e_L}(t)/G_{e_H}(t))^N}$$

and therefore $t_1^* < t_0^*$. Furthermore, we clearly have $t_0^*,\ t_1^* \to 0$ as $N \to \infty$.

Now, because $\frac{G_{e_H}(t)}{G_{e_L}(t)}$ is monotone increasing in t, the threshold t_B^* is monotone increasing in C_H and $t_B^* \downarrow 0$ when $C_H \downarrow C_L$, and t_B^* converges to $+\infty$ when C_H increases. Therefore, there exist thresholds $C_{B,1}^* < C_{B,2}^*$ such that $t_B^* = t_1^*$ when $C_H = C_{B,1}^*$ and $t_B^* = t_0^*$ when $C_H = C_{B,2}^*$.

Because the functions

$$\phi_i(t) \equiv \frac{(e^{(\gamma-r)t}-i)}{1-(G_{e_L}(t)/G_{e_H}(t))^N}$$

are monotone decreasing (increasing) for $t \le t_i^*$ ($t \ge t_i^*$), $i = 0, 1$, we get that, for $C_H < C_{B,1}^*$,

$$\begin{aligned} U_S^0 + C_H + \frac{(C_H - C_L)(e^{(\gamma-r)t_1^*}-1)}{1-(G_{e_L}(t_1^*)/G_{e_H}(t_1^*))^N} &\le U_S^0 + C_H + \frac{(C_H - C_L)(e^{(\gamma-r)t_B^*}-1)}{1-(G_{e_L}(t_B^*)/G_{e_H}(t_B^*))^N} \\ = \frac{e^{(\gamma-r)t_B^*}(C_H - C_L)}{1-(G_{e_L}(t_B^*)/G_{e_H}(t_B^*))^N}, \end{aligned} \tag{65}$$

hence the contract corresponding to t_1^* is optimal.

When $C_{B,1}^* < C_H < C_{B,2}^*$, we have $t_1^* < t_B^* < t_0^*$ and therefore the quantity (71) is equal to (70) and the minimum is attained when $y_0 = 0$ and $t = t_B^*$. Finally, when $C_H > C_{B,2}^*$, we have $t_B^* > t_0^* > t_1^*$ and therefore the minimum in (71) is still attained at $t = t_B^*$, whereas the maximum in (70) is attained at $t = t_0^*$ and is therefore strictly smaller.

(2) Suppose now that the intermediary has full bargaining power. Then, the same arguments used in case 1 imply that the maximization problem for the intermediary

takes the form

$$\max\left\{y_0 \;+\; \frac{(C_H - C_L)}{1-(G_{e_L}(t)/G_{e_H}(t))^N} \;:\; y_0 + \frac{e^{(\gamma-r)t}\,(C_H - C_L)}{1-(G_{e_L}(t)/G_{e_H}(t))^N} \;\leq\; U_B(\{\delta_n\}, e_H) - U_B^0\right\}$$

Thus, we are maximizing over the set

$$\left\{t:\; \frac{e^{(\gamma-r)t}\,(C_H - C_L)}{1-(G_{e_L}(t)/G_{e_H}(t))^N} \;\leq\; U_B(\{\delta_n\}, e_H) - U_B^0\right\},$$

which coincides with a segment $[\underline{t}_S, \bar{t}_S]$ satisfying $\underline{t}_S \leq t_0^* \leq \bar{t}_S$. The optimal y_0 is always given by

$$y_0 \;=\; -\frac{e^{(\gamma-r)t}\,(C_H - C_L)}{1-(G_{e_L}(t)/G_{e_H}(t))^N} \;+\; U_B(\{\delta_n\}, e_H) - U_B^0\,.$$

Hence, the maximization problem takes the form

$$\max_{t\in[\underline{t}_S, \bar{t}_S]}\left(U_B(\{\delta_n\}, e_H) - U_B^0 \;-\; \frac{(e^{(\gamma-r)t}-1)\,(C_H - C_L)}{1-(G_{e_L}(t)/G_{e_H}(t))^N}\right).$$

Because $t_1^* < t_0^* < \bar{t}_S$, the maximum is attained at $\max\{t_1^*, \underline{t}_S\}$. Clearly, $\underline{t}_S$ is the minimal solution to

$$\frac{e^{(\gamma-r)t}}{1-(G_{e_L}(t)/G_{e_H}(t))^N} \;=\; \frac{U_B(\{\delta_n\}, e_H) - U_B^0}{C_H - C_L}$$

and is therefore increasing in C_H. Furthermore, $\underline{t}_S \downarrow 0$ as $C_H \downarrow C_L$ and $\underline{t}_S \to 0$ as $N \to \infty$. Therefore, there exists a threshold C_S^* such that $t_1^* > \underline{t}_S$ if and only if $C_H < C_S^*$, and the required assertion follows. ■

Proof of Proposition 4..9. By definition, t_1^* solves

$$\begin{aligned} 0 \;&=\; \frac{d}{dt}\frac{e^{(\gamma-r)t}-1}{1-(G_{e_L}(t)/G_{e_H}(t))^N} \\ &=\; \frac{(\gamma-r)e^{(\gamma-r)t}(1-(G_{e_L}(t)/G_{e_H}(t))^N) - (e^{(\gamma-r)t}-1)N(G_{e_L}(t)/G_{e_H}(t))^N\,(h_{e_L}(t)-h_{e_H}(t))}{(1-(G_{e_L}(t)/G_{e_H}(t))^N)^2}\,. \end{aligned} \tag{66}$$

Because t_1^* is a local minimum, this derivative changes sign from negative to positive at t_1^*. Therefore, the same is true for the function

$$\phi(t) \;\equiv\; 1 - \frac{N(h_{e_L}(t) - h_{e_H}(t))(1 - e^{-(\gamma - r)t})}{(\gamma - r)((G_{e_H}(t)/G_{e_L}(t))^N - 1)},$$

and hence $\phi'(t_1^*) > 0$.[41]

A direct calculation shows that, for any $A > 1$, the function $(A^y - 1)/y$ is monotone increasing in y, whereas $(1 - A^{-y})/y$ is monotone decreasing in y. Therefore,

$$\frac{\partial \phi}{\partial N} > 0 \;,\; \frac{\partial \phi}{\partial(\gamma - r)} \;>\; 0\,.$$

Differentiating the identity $\phi(t_1^*) = 0$, we get

$$\frac{\partial t_1^*}{\partial N} \;=\; -\frac{\partial \phi/\partial N}{\partial \phi/\partial t_1^*}|_{t=t_1^*} \;<\; 0$$

and the same is true for $\gamma - r$.

Similar arguments imply that

- t_B^* is monotone decreasing in $N,\ C_L$ and U_S^0 and is monotone increasing in C_H;
- t_0^* is monotone decreasing in $\gamma - r$ and N. In particular, the maturity of the optimal contract is then also decreasing in N and $\gamma - r$;
- $\underline{t}_S$ is monotone decreasing in N and is increasing in $U_B^0, C_H - C_L$ and $\gamma - r$.

Finally, $\underline{t}_S < t_0^*$ by the arguments given.

The proof is complete. ■

[41] For simplicity we assume that the inequality is strict. This is true for generic parameter values. The general case can be considered by a small modification of the argument.

Proposition D.2 *Under the hypothesis of Theorem 4..5, the total second best surplus* SB_H *for high effort satisfies the following:*

- *In the monopolistic case, the total surplus for the investor is given by*

$$\mathrm{SB}_H(N) \;=\; \begin{cases} \mathrm{FB}_H - (C_H - C_L)\phi_1(t_1^*)\;, & C_H < C_{B,1}^* \\ \mathrm{FB}_H - (e^{(\gamma-r)t_B^*} - 1)(U_S^0 + C_H)\;, & C_H \in [C_{B,1}^*, C_{B,2}^*) \\ \mathrm{FB}_H - ((C_H - C_L)\phi_0(t_0^*) - (U_S^0 + C_H))\;, & C_H > C_{B,2}^*. \end{cases}$$

- *In the competitive case, the total surplus for the intermediary is given by*

$$\begin{aligned} &\mathrm{SB}_H(N) \\ &= \begin{cases} \mathrm{FB}_H - (C_H - C_L)\phi_1(t_1^*)\;, & C_H < C_S^* \\ \mathrm{FB}_H \;-\; \Big((1 - e^{-(\gamma-r)t^*})(U_B(\{\delta_n\}, e_H) - U_B^0) - (U_S^0 + C_H)\Big)\;, & C_H \geq C_S^*\,. \end{cases} \end{aligned} \tag{67}$$

In particular, for $C < \min\{C_{B,1}^*, C_S^*\}$*, the two surpluses are identical.*

Proof of Proposition 5..2. We only consider the case $C_H < C^*$. Other cases are similar.

Because $G_L(t)/G_H(t) < 1$, we have that $(G_L(t)/G_H(t))^N$ is monotone decreasing in N. Consequently,

$$\phi_i(t) \;=\; \frac{e^{(\gamma-r)t} - i}{1 - (G_L(t)/G_H(t))^N}$$

is also decreasing in N.

Now, fix an $\varepsilon > 0$ and let $\tau > 0$ be such that $e^{(\gamma-r)\tau} - 1 < \varepsilon/2$. Then, there exists an $N(\tau)$ such that $(G_L(t)/G_H(t))^N < 0.5$ for all $N > N(\tau)$. Therefore, for all $N > N(\tau)$, we have

$$\min_{t>0} \phi_1(t) \;\leq\; \phi_1(\tau) < \frac{\varepsilon/2}{0.5} \;=\; \varepsilon,$$

and the required assertion follows. ■

It follows directly from Proposition B.2 that the following is true.

Proposition D.3 *In the Black and Cox setting, suppose that the effort is binary, higher effort reduces default risk, but* $m_H/\sigma_H < m_L/\sigma_L$. *Then,*

$$\max_{\tau_{[1,n]} \leq t} x_n(t; \tau_{[1,n]}) \;=\; x_n(t; (t, \cdots, t))$$

and therefore $x_n(t; \tau_{[1,n]})$ *is identically zero for* $t \geq \bar{t}_n$, *where*

$$\bar{t}_n \;=\; \max\left\{ t : \; e^{(r-\gamma)t} \left(\mu_1 \left(1 - \frac{(p_{e_L}(t))^n (G_{e_L}(t))^{N-n}}{(p_{e_H}(t))^n (G_{e_H}(t))^{N-n}} \right) + \mu_2 \right) \;\geq\; 1 \right\} .$$

Proposition D.4 (The threshold model) *Suppose that*

$$p_{e_L}(t) \;=\; \mathbf{1}_{t \geq s} \lambda_H \, e^{-\lambda_H (t-s)} \;,\; p_{e_H}(t) \;=\; \mathbf{1}_{t \geq s} \lambda_L \, e^{-\lambda_L (t-s)} .$$

Let also

$$t_B^* \;=\; \frac{1}{(\lambda_H - \lambda_L) N} \log\left(\frac{U_S^0 + C_H}{U_S^0 + C_L} \right) \;,\; t_0^* \;=\; \frac{1}{(\lambda_H - \lambda_L) N} \log\left(1 + \frac{(\lambda_H - \lambda_L) N}{\gamma - r} \right)$$

Then, the following is true:

and the maturity of the optimal contract is $\min\{t_B^*, t_0^*\}$ *in the monopolistic case and* $\underline{t}_S$ *in the competitive case.*

Proof of Proposition 4..6. It follows directly from the above that the required assertion holds if and only if $t_1^* > 0$. The claim thus immediately follows from $\phi_1(0) = +\infty$. ■

Proof of Proposition 4..7. First, it follows by direct calculation that the distributions are 2-regular and the local maxima can only happen at $t = 0$ and a single

positive t^*.

(1) Suppose that the investor has full bargaining power. Then, the investor's maximization problem takes the form

$$\min_{t,y_0}\left\{y_0+\frac{e^{(\gamma-r)t}\,(C_H-C_L)}{1-(G_{e_L}(t-s)/G_{e_H}(t-s))^N}\ :\ y_0\geq 0,\ y_0+\frac{(C_H-C_L)}{1-(G_{e_L}(t-s)/G_{e_H}(t-s))^N}\geq U_S^0+C_H\right\}$$

Because $p_{e_L}\prec_{hr} p_{e_H}$, the quotient $\frac{G_{e_H}(t)}{G_{e_L}(t)}$ is monotone increasing in t; therefore, a direct calculation shows that

$$U_S^0+C_H\ \leq\ \frac{C_H-C_L}{1-(G_{e_L}(t^*-s)/G_{e_H}(t^*-s))^N)} \tag{68}$$

is equivalent to

$$t^*\leq s+t_B^*\,,$$

where $t_B^*\ =\ t_B^*$ is the unique solution to

$$\frac{G_{e_H}(t_B^*)}{G_{e_L}(t_B^*)}\ =\ \left(\frac{U_S^0+C_H}{U_S^0+C_L}\right)^{1/N}\,. \tag{69}$$

Note that this unique solution only exists if

$$\frac{P_{e_H}^\infty}{P_{e_L}^\infty}\ >\ \left(\frac{U_S^0+C_H}{U_S^0+C_L}\right)^{1/N}\,.$$

Otherwise, we set $t_B^*=+\infty$. Clearly, $t_B^*\to 0$ as $N\to\infty$.

Thus, if $t<s+t_B^*$, the optimal choice is clearly $y_0=0$ and the cost minimization problem takes the form

$$\min_{t\leq t_B^*}\frac{e^{(\gamma-r)t}\,(C_H-C_L)}{1-(G_{e_L}(t-s)/G_{e_H}(t-s))^N}\,, \tag{70}$$

and the minimum is clearly attained at $s + \min\{t_B^*, t_0^*\}$.

If $t > t_B^* + s$, y_0 needs to be positive to satisfy the IR constraint of the seller, and the optimal choice is clearly

$$y_0 \;=\; U_S^0 + C_H - \frac{(C_H - C_L)}{1 - (G_{e_L}(t-s)/G_{e_H}(t-s))^N}\,.$$

Therefore, the cost minimization problem takes the form

$$\min_{t \geq t_B^* + s} \left(U_S^0 + C_H + \frac{(C_H - C_L)(e^{(\gamma - r)t} - 1)}{1 - (G_{e_L}(t-s)/G_{e_H}(t-s))^N} \right), \tag{71}$$

and the minimum is clearly attained at $s + \max\{t_B^*, t_1^*(s)\}$, where $t_1^*(s)$ minimizes

$$\frac{(C_H - C_L)(e^{(\gamma - r)s} e^{(\gamma - r)t} - 1)}{1 - (G_{e_L}(t)/G_{e_H}(t))^N}$$

over $t \geq 0$.

Lemma D.5 *We have $t_1^*(0) = 0$ and $t_1^*(s)$ is monotone increasing in s and converges to t_0^* as $s \to \infty$. Thus, there exists a critical $\bar{s} > 0$ such that $t_1^*(\bar{s}) = t_B^*$ and $t_1^*(s) > t_B^*$ for all $s > \bar{s}$.*

Proof. First, a direct calculation shows that the function $(e^a - 1)/a$ is monotone increasing in a. Indeed, differentiating, we get that this is equivalent to $e^a(a-1) + 1 > 0$ which is in turn equivalent to $e^{-a} > 1 - a$, which follows from the convexity of e^{-a}. This immediately implies that

$$\frac{e^a - 1}{a} \;>\; \frac{e^{-b} - 1}{-b} \;=\; \frac{1 - e^{-b}}{b} \;\Rightarrow\; \frac{e^a - 1}{1 - e^{-b}} > \frac{a}{b}$$

for any $a, b > 0$. Thus,

$$\frac{e^{(\gamma-r)t}-1}{1-e^{-(\lambda_H-\lambda_L)Nt}} \;>\; \frac{\gamma-r}{(\lambda_H-\lambda_L)N} \;=\; \lim_{t\downarrow 0}\frac{e^{(\gamma-r)t}-1}{1-e^{-(\lambda_H-\lambda_L)Nt}},$$

and hence $t_1^*(0) = 0$. Monotonicity in s follows by the same arguments as above. Finally, convergence to t_0^* follows because

$$t_1^* = \arg\min \frac{(C_H - C_L)(e^{(\gamma-r)t} - e^{-(\gamma-r)s})}{1-(G_{e_L}(t)/G_{e_H}(t))^N}$$

and this function converges to $\phi_0(t)$ as $s \to \infty$. ■

The minimal cost is then given by the minimum of the two quantities above.

Furthermore, a direct calculation shows that

$$\frac{d}{dt}\frac{(C_H - C_L)(e^{(\gamma-r)(t+s)}-1)}{1-(G_{e_L}(t)/G_{e_H}(t))^N} \;>\; \frac{d}{dt}\frac{e^{(\gamma-r)(t+s)}\,(C_H - C_L)}{1-(G_{e_L}(t)/G_{e_H}(t))^N}$$

and therefore $t_1^* < t_0^*$. Furthermore, we clearly have $t_0^*,\ t_1^* \;\to\; 0$ as $N \to \infty$.

Now, because $\frac{G_{e_H}(t)}{G_{e_L}(t)}$ is monotone increasing in t, the threshold t_B^* is monotone increasing in C_H and $t_B^* \downarrow 0$ when $C_H \downarrow C_L$, and t_B^* converges to $+\infty$ when C_H increases. Therefore, there exist thresholds $C_{B,1}^*(s) < C_{B,2}^*$ such that $t_B^* = t_1^*$ when $C_H = C_{B,1}^*(s)$ and $t_B^* = t_0^*$ when $C_H = C_{B,2}^*$.

Because the functions

$$\phi_i(t) \;\equiv\; \frac{(e^{(\gamma-r)(t+s)} - i)}{1-(G_{e_L}(t)/G_{e_H}(t))^N}$$

are monotone decreasing (increasing) for $t \leq t_i^*$ ($t \geq t_i^*$), $i = 0, 1$, we get that, for $C_H < C_{B,1}^*(s)$,

$$\begin{aligned} U_S^0 + C_H + \frac{(C_H - C_L)(e^{(\gamma-r)(t_1^*+s)} - 1)}{1 - (G_{e_L}(t_1^*)/G_{e_H}(t_1^*))^N} &\leq U_S^0 + C_H + \frac{(C_H - C_L)(e^{(\gamma-r)(t_B^*+s)} - 1)}{1 - (G_{e_L}(t_B^*)/G_{e_H}(t_B^*))^N} \\ &= \frac{(C_H - C_L)}{1 - (G_{e_L}(t_B^*)/G_{e_H}(t_B^*))^N} + \frac{(C_H - C_L)(e^{(\gamma-r)(t_B^*+s)} - 1)}{1 - (G_{e_L}(t_B^*)/G_{e_H}(t_B^*))^N} \\ &= \frac{e^{(\gamma-r)(t_B^*+s)}\,(C_H - C_L)}{1 - (G_{e_L}(t_B^*)/G_{e_H}(t_B^*))^N}\,, \end{aligned} \tag{72}$$

hence the contract corresponding to $t_1^*(s)$ is optimal and $y_0 > 0$.

When $C_{B,1}^* < C_H < C_{B,2}^*$, we have $t_1^*(s) < t_B^* < t_0^*$ and therefore the quantity (71) is equal to (70) and the minimum is attained when $y_0 = 0$ and $t = t_B^* + s$. Finally, when $C_H > C_{B,2}^*$, we have $t_B^* > t_0^* > t_1^*$ and therefore the minimum in (71) is still attained at $t = t_B^* + s$, whereas the maximum in (70) is attained at $t = t_0^* + s$ and is therefore strictly smaller.

(2) Suppose now that the intermediary has full bargaining power. Then, the same arguments used in case 1 imply that the maximization problem for the intermediary takes the form

$$\begin{aligned} \max \Big\{ y_0 \;+\; & \frac{(C_H - C_L)}{1 - (G_{e_L}(t-s)/G_{e_H}(t-s))^N} \;: \\ & y_0 + \frac{e^{(\gamma-r)t}\,(C_H - C_L)}{1 - (G_{e_L}(t-s)/G_{e_H}(t-s))^N} \;\leq\; U_B(\{\delta_n\}, e_H) - U_B^0 \Big\} \end{aligned} \tag{73}$$

Thus, we are maximizing over the set

$$\left\{ t : \; \frac{e^{(\gamma-r)t}\,(C_H - C_L)}{1 - (G_{e_L}(t-s)/G_{e_H}(t-s))^N} \;\leq\; U_B(\{\delta_n\}, e_H) - U_B^0 \right\},$$

which coincides with a segment $[\underline{t}_S(s), \bar{t}_S(s)]$ satisfying $\underline{t}_S \leq t_0^* + s \leq \bar{t}_S$. The optimal y_0 is always given by

$$y_0 \;=\; -\frac{e^{(\gamma-r)t}\,(C_H - C_L)}{1-(G_{e_L}(t-s)/G_{e_H}(t-s))^N} \;+\; U_B(\{\delta_n\}, e_H) - U_B^0\,.$$

Hence, the maximization problem takes the form

$$\max_{t\in[\underline{t}_S,\bar{t}_S]} \left(U_B(\{\delta_n\}, e_H) - U_B^0 \;-\; \frac{(e^{(\gamma-r)t}-1)\,(C_H - C_L)}{1-(G_{e_L}(t-s)/G_{e_H}(t-s))^N} \right)\,.$$

Because $t_1^*(s) + s < t_0^*(s) + s < \bar{t}_S$, the maximum is attained at $\max\{s + t_1^*(s), \underline{t}_S(s)\}$. Clearly, $\underline{t}_S - s$ is the minimal solution to

$$\frac{e^{(\gamma-r)t}}{1-(G_{e_L}(t)/G_{e_H}(t))^N} \;=\; e^{-(\gamma-r)s}\frac{U_B(\{\delta_n\}, e_H) - U_B^0}{C_H - C_L}$$

and is therefore $\underline{t}_S - s$ is monotone decreasing in s, and the existence of the threshold $\bar{s}$ follows.

■

Proof of Proposition 4..8. As above, we need to show that $t_1^* > 0$. To this end, it suffices to show that $\phi_1(t) < \phi_1^(0)$ for sufficiently small t. This follows by direct calculation from the Taylor formula and the l'Hopital rule. ■

E Securitization and equilibrium effort level in the risk neutral case

The following claim follows by direct calculation

Lemma E.1 *In the risk neutral case, we have*

$$U_B(\{\delta_n\}, e_j) \;=\; N\,u\,r^{-1}\left(1-(1-R/u)\int_0^\infty e^{-rt}p_{e_j}(t)dt\right)$$

and

$$U_S(\{\delta_n\}, e_j) \;=\; N\,u\,\gamma^{-1}\left(1-(1-R/u)\int_0^\infty e^{-\gamma t}p_{e_j}(t)dt\right) \;-\; C_j.$$

Thus, without securitization, the seller chooses high effort if and only if

$$C_H - C_L \;\leq\; \gamma^{-1}(u-R)N\int_0^\infty e^{-\gamma t}(p_{e_L}(t)-p_{e_H}(t))\,dt \;=\; \overline{C}. \tag{74}$$

Lemma E.2 *In the risk neutral case, the optimal contract implementing low effort simply makes a single payment at time* 0*. This payment is equal to*

$$U_S^0 + C_L$$

when the investor has full bargaining power, and to

$$U_B(\{\delta_n\}, e_L) - U_B^0$$

when the intermediary has full bargaining power.

Proof. Because $e^{-\gamma t} < e^{-rt}$ for all $t > 0$, we have

$$E\left[\int_0^\infty e^{-\gamma t}dX_t|e_L\right] \;\leq\; E\left[\int_0^\infty e^{-rt}dX_t|e_L\right]$$

and the inequality is strict if dX_t is positive for $t > 0$ with positive probability. In the case where the investor has full bargaining power, the intermediary's PC constraint gives

$$\min_{dX \geq 0} E\left[\int_0^\infty e^{-rt} dX_t | e_L\right] \;\geq\; U_S^0,$$

and in the case where the intermediary has full bargaining power, the investor's PC constraint gives

$$\max_{dX \geq 0} E\left[\int_0^\infty e^{-\gamma t} dX_t | e_L\right] \;\leq\; U_B(\{\delta_n\}, e_L) - U_B^0\,.$$

■