The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXII, NO. 5 • OCTOBER 2017

The Downside of Asset Screening for Market Liquidity

VICTORIA VANASCO*

ABSTRACT

This paper explores the tension between asset quality and market liquidity. I model an originator who screens assets whose cash flows are later sold in secondary markets. Screening improves asset quality but gives rise to asymmetric information, hindering trade of the asset cash flows. In the optimal mechanism (second-best), costly retention of cash flows is essential to implement asset screening. Market allocations can feature too much or too little screening relative to second-best, where too much screening generates inefficiently illiquid markets. Furthermore, the economy is prone to multiple equilibria. The optimal mechanism is decentralized with two tools: retention rules and transfers.

SECONDARY MARKETS FOR ASSETS PLAY an important role in providing lending capacity to the financial industry and the real economy, by allowing financial institutions to raise funds through asset sales. In 2007 more than 25% of outstanding consumer credit in the United States was financed through the securitization of consumer loans.[1] Following the financial crash of 2008, however, issuance of securitized assets collapsed, adversely affecting financial institutions and, in turn, firms and consumers' access to credit (Chodorow-Reich (2014), Mondragon (2015)). In response, policy makers geared their efforts

*Victoria Vanasco is from the Graduate School of Business, Stanford University. I thank Bruno Biais (Editor), the Associate Editor, and two anonymous referees. I am also thankful to Vladimir Asriyan, Mitch Berlin (discussant), Peter DeMarzo, Sebastian Di Tella, Darrel Duffie, Dana Foarta, Alex Frankel, William Fuchs, Pierre-Olivier Gourinchas, Brett Green, Laurie Hodrick, Amir Kermani, Arvind Krishnamurthy, Ulrike Malmendier, Takeshi Murooka, Aniko Öery, Christine Parlour, Giorgia Piacentino, Demian Pouzo, David Romer, Nancy Wallace, and Yao Zeng for insightful discussions and suggestions; seminar participants at UC Berkeley Department of Economics, Haas School of Business, GSB at Stanford University, The Fuqua School of Business, University of Maryland, Universita Bocconi, EIEF Rome, CREI and Universitat Pompeu Fabra, University of Toronto, HEC Paris, Columbia Business School, Wharton School of Business, the New York and Richmond Federal Reserves, the Federal Reserve Board, and the European Central Bank; and conference participants at the FTG Meeting at UCLA (2014), the Econometric Society Meetings in Madrid (2014), FIRS (2015), and UTDT Economics Conference (2015). I acknowledge financial support from The Clayman Institute for Gender Research and I declare that I have no relevant or material financial interests related to the research in this paper. All errors are my own.

[1] Securitization is the practice of creating and selling securities whose payoffs are derived from and collateralized by a specified pool of underlying assets. Some common examples include (commercial) mortgage-backed securities ((C)MBS), collaterilized loan obligations (CLO), and asset-backed securities (ABS).

DOI: 10.1111/jofi.12519

toward reviving these markets.[2] Two key market frictions have been brought to light. First, the ability to sell loan cash flows in secondary markets (directly or through securitization) has been associated with a decline in lending standards (see Berndt and Gupta (2008) for the syndicated loan market, and DellAriccia, Igan, and Laeven (2012), Elul (2016), Jaffee et al. (2009), Keys et al. (2010), and Mian and Sufi (2009) for the mortgage market). Second, originators have used private information about loan quality when choosing which loans to securitize (Downing, Jaffee, and Wallace (2008), Calem, Henderson, and Liles (2011), Agarwal, Chang, and Yavas (2012), Jiang, Nelson, and Vytlacil (2014)). There is a problem of incentives at the asset origination stage followed by a problem of asymmetric information that may limit the ability to trade these assets (i.e., that reduces asset liquidity). This raises two questions: how should originators be incentivized to screen asset quality while preserving liquidity in secondary markets for these assets, and is there a need for policy intervention?

To address these questions, I develop a theoretical framework to study the trade-off between an asset's quality and liquidity. In the model, an originator can exert costly effort to improve the quality of an asset, the cash flows of which can be sold in secondary markets. The key feature of the model is that both the level of the screening effort and the quality of the originated asset are the originator's private information. This setup aims to capture the fact that not all of the information that is acquired while screening or monitoring can be easily conveyed to outside investors.[3] Asset screening may therefore expose the originator to illiquid secondary markets for the asset due to her private information. This is costly when there are gains from trade with outside investors. In this setting, I explore the resulting tension between productive efficiency—the quality of originated assets—and allocative efficiency—the final allocation of asset cash flows.

This framework is general and can be applied to many economic settings. The premise is that an agent's hidden action determines the distribution of output and gives the agent superior information about the realization of this output. This superior information can, in turn, affect the liquidity of claims on this output. All originators of informationally sensitive financial assets face this trade-off. This tension is also present in nonfinancial settings where an agent, such as a venture capitalist (VC) or a CEO with stock options, can exert effort to improve a firm's cash flows. When monitoring generates private information, too much monitoring may impact the liquidity of the agent's stocks (Faure-Grimaud and Gromb (2004)) or exit options (Aghion, Bolton, and Tirole (2004)). Other examples include the relationship between firm control and its impact on stock market trading (Coffee (1991), Bhide (1993), Dewatripont and Tirole (1994), Maug (1998), Medrano and Vives (2004)).

[2] The 2010 report of the Financial Stability Board to G20 leaders stated "re-establishing securitization on a sound basis remains a priority in order to support provision of credit to the real economy and improve banks' access to funding" (FSB (2010)).

[3] See Qian, Strahan, and Yang (2015) for evidence on how incentives and communication costs affect information production in banking.

Turning back to the setting of interest here, the model has three periods and features a manager and market investors. The manager can finance and manage one risky project that pays off in the final period. In the first period, she exerts costly effort to screen projects to increase the likelihood of finding a "good" project. If a "good" project is found, it is financed; if not, an average project from the pool of potential projects is financed. Both the screening effort and the quality of the financed project are the manager's private information. In the intermediate period, the manager can exploit gains from trade with investors, who are more patient, by selling securities that are backed by the project's cash flows. The cash flows that are not sold are retained by the manager. In the first-best of this economy, the manager (i) chooses screening effort in the first period such that the social marginal benefit of asset screening equals its social marginal cost—full productive efficiency, and (ii) sells all of her asset cash flows in the intermediate period to investors—full allocative efficiency.

Full productive and allocative efficiency cannot be achieved in the presence of information frictions. I thus characterize the optimal mechanism that maximizes ex ante efficiency (second-best). I then describe the equilibrium allocations and study the conditions under which they differ from the second-best allocations. Under the optimal mechanism, retention of cash flows is essential to implement positive effort. Ex ante efficiency is maximized with differential retention levels—managers with poor-quality assets should not retain as much as those with good-quality assets. In contrast, equilibrium allocations can feature either too little or too much screening effort relative to the second-best. While the presence of asymmetric information is essential to sustaining the equilibrium with positive effort, too much effort may generate inefficiently illiquid assets.

The predictions of the model shed light on the observed booms and busts in origination and securitization of some asset classes around the 2008 to 2009 financial crisis. First, as gains from trade increase, which can be interpreted as an increase in liquidity needs, trade in secondary markets increases while the quality of underlying asset cash flows decreases. Second, when gains from trade are large enough, investors' beliefs about the manager's screening effort become self-fulfilling, which increases market fragility. A second equilibrium with no asset screening and no cash flow retention arises. This behavior is in line with the observed trend in the United States of certain securitized assets, such as nonagency MBS, that featured a boom in the years leading up to the crisis that was accompanied by a decrease in the quality of the underlying loans (Mian and Sufi (2009)).

Under the optimal mechanism, while retention of cash flows is essential to implement positive screening effort, it reduces gains from trade. As a result, effort is optimally chosen to trade off the social benefit of asset screening with its social cost, which includes the indirect cost of the cash flow retention required to implement it. This indirect cost generates a wedge between the first-best and the second-best allocations. The notion that cash flow retention incentivizes asset screening is consistent with evidence in Ashcraft, Gooriah, and Kermani (2014) that an increase in the amount of cash flow retention is correlated with better security performance in the conduit CMBS market.[4] Cash flow retention thus plays a dual role in the implementation of screening effort. First, it directly exposes the manager to her choice of screening effort. Second, and more importantly, it is needed to identify (i.e., to separate) the quality of the manager's asset, which is *good* if a "good" project was financed and *bad* if an "average" project was financed.

Requiring that all originators (all manager types) have the same retention level is inefficient in the presence of adverse selection. It is not the best way to provide incentives to screen assets, and it decreases allocative efficiency. Retention is less costly for good managers and as a result efficiency is improved when the retention of the good manager increases while that of the bad one decreases. Differential retention levels are able to implement a given effort level with less overall retention. Cross-subsidization also increases efficiency: all secondary market surplus should be transferred to the good manager, subject to the bad one not mimicking. Such transfers allow the mechanism to separate types for any positive retention level. Thus, transfers across managers impact ex ante efficiency by influencing both retention decisions and screening effort choices.

Market equilibrium allocations can feature too much or too little screening effort relative to the second-best. These allocations differ from those obtained under the optimal mechanism due to lack of commitment in the first period to choices made after the level of effort is chosen. This introduces two externalities relative to the second-best. First, the manager does not internalize the effect of her effort choice on the quality of cash flows sold to investors—*effort externality*. Second, the manager does not internalize the effect of cash flow retention on her ex ante value, since the former is chosen ex post to signal quality—*retention externality*. Which externality dominates determines whether the market features inefficiently low or high cash flow retention and screening effort. When both the manager and investors can commit to the retention levels and to the prices chosen before projects are screened, the market equilibrium implements the second-best allocations.

One of the insights of the paper is that the presence of asymmetric information alleviates the problem of incentives; that is, adverse selection is essential to implement positive effort in markets with no commitment. To illustrate this point, consider an alternative setting in which the manager exerts hidden screening effort but does not have private information about the quality of her asset. I show that in this case the market equilibrium features neither asset screening nor retention. Furthermore, the optimal mechanism achieves higher ex ante efficiency when the manager has private information about her asset quality than when she does not. This is because the mechanism can extract this information and use it to improve both productive and allocative efficiency.

[4] Commercial MBS are generated by "conduit" lenders, who originate commercial mortgages with the aim of arranging them into pools and selling standardized sections of these pools to investors.

After characterizing the market allocations, I study the policy implications of the model. I show that the optimal mechanism can be decentralized in markets with no commitment with two policy tools: differential retention levels and transfers across markets for different securities. In particular, I find that lump-sum taxes on the issuance of junior claims together with subsidies for the issuance of senior claims should accompany policies that impose retention levels on originators. Policy plays a dual role in this environment. First, it improves ex ante efficiency by affecting both the screening effort and the level of trade in secondary markets. Second, and equally important, it provides market stability, as it eliminates the multiplicity of equilibria that may arise and it ensures the existence of a unique equilibrium with positive screening. In the presence of two-sided commitment (by the manager and by investors), there is no need for policy intervention. However, since this requires commitment to contracts that are not necessarily renegotiation-proof, in reality, many markets may be represented by the no-commitment assumption.[5]

The policy implications of this paper apply more directly to financial institutions that originate financial assets than to active firm management and monitoring, since the ability to commit to predesigned contracts is likely to be different in these settings. For example, while contracts on VC exit strategies and option exercise are actually written out in practice, agreements between asset originators and final investors are more subtle and implicit (if present at all). Evidence suggesting the absence of commitment in markets for securitized assets is presented by Ashcraft, Gooriah, and Kermani (2014), who find that the price at issuance of senior securities in the CMBS market is not correlated with the amount of retention, that is, cash flow retention cannot be predicted by prices at origination. However, the characterization of the optimal mechanism does apply to both settings and can be used to think about the optimal design of contracts.

This paper contributes to the discussion on how to regulate markets for securitized assets. Policy makers worldwide have agreed on the "skin in the game" rule, which requires that originators and securitizers retain a risk exposure to the underlying assets. My model rationalizes this type of intervention, but suggests that demanding the same retention level of all originators may impose excessive retention costs. In addition, regulators should not only focus on cash flow retention, but also on how issuers are compensated in the market since cross-subsidization further improves both liquidity and incentives to screen.

Several papers highlight the trade-off between incentives to originate good assets and secondary market liquidity of the asset cash flows. The idea that secondary market liquidity affects incentives to screen asset quality is explored by Parlour and Plantin (2008), Plantin (2011), Malherbe (2012), and Chemla and Hennessy (2014). My paper contributes to this line of research by modeling an environment in which screening increases the likelihood of originating a good asset (which is desirable), but also affects the distribution of the asset

[5] See Hartman-Glaser (2017) and Chari, Shourideh, and Zetlin-Jones (2014) for the effect of reputational concerns in markets with adverse selection.

cash flows and worsens adverse selection (which is detrimental). This dual role of screening has important implications for equilibrium outcomes and the resulting inefficiencies. For example, market allocations may feature excessive screening effort relative to second-best allocations. This inefficiency does not arise in the above-mentioned papers, where liquidity is independent of screening effort. In addition, my model features multiplicity of equilibria, which to my knowledge is novel in this type of environment.

My work also relates to the extensive literature on security design and loan sales in the presence of adverse selection started by Leland and Pyle (1977), Myers and Majluf (1984), and Gorton and Pennacchi (1995). While I do not solve a full blown security design problem, the forces that determine the cash flows sold and prices paid in secondary markets in this paper are as those described in DeMarzo and Duffie (1999), DeMarzo (2005), and Biais and Mariotti (2005), where the latter also explore the optimal mechanism. I contribute to this literature by endogeneizing the decision to originate assets in order to study the trade-off between asset quality and market liquidity. My paper is therefore also related to the literature on security design in the presence of moral hazard (Innes (1990), Fender and Mitchell (2009), Hartman-Glaser, Piskorski, and Tchistyi (2012), Yang and Zeng (2014), Hébert (2015)).

The remainder of the paper is organized as follows. Section I introduces the model setup and characterizes the first-best allocations. Section II presents the optimal mechanism: the second-best. Section III describes the equilibrium allocations and Section IV relates them to the second-best. Section V concludes. All proofs are relegated to the Appendix.

I. The Model

A. Screening and Investment Technology

The model has three periods, indexed by $t \in \{0, 1, 2\}$. There is an originator, who I refer to as a manager, and a market of potential investors. All agents are risk-neutral with utility functions $V_0(c_1, c_2) = \theta c_1 + c_2$ and $V_0^i(c_1^i, c_2^i) = c_1^i + c_2^i$, where V_0 (V_0^i) denotes the utility in $t = 0$ and c_t (c_t^i) the cash flows at time t of the manager (investor i), and $\theta > 1$ denotes the manager's marginal value of funds in $t = 1$. The different valuation of $t = 1$ cash flows generates gains from trade between the manager and investors.[6] The manager can finance and manage one risky project chosen in $t = 0$. When the manager finances a project, she originates an asset (e.g., a loan).

Project Screening and Investment. There is a pool of risky projects available to the manager. Each project pays cash flows X in $t = 2$, where $X = X_H$ with probability $\pi \in (0, 1)$ and $X = X_L$ with probability $1 - \pi$. I suppose that $X_H > X_L$

[6] Modeling gains from trade as $\theta > 1$ is standard in the literature to capture gains from selling assets not explicitly modeled in this paper (see Holmström and Tirole (2011)). There are many reasons why a manager might need to sell her assets: capital constraints, new investment opportunities, risk-sharing, creation of bankruptcy remote instruments, etc. For an interpretation of the gains from trade in the active monitor setting, see Aghion, Bolton, and Tirole (2004).

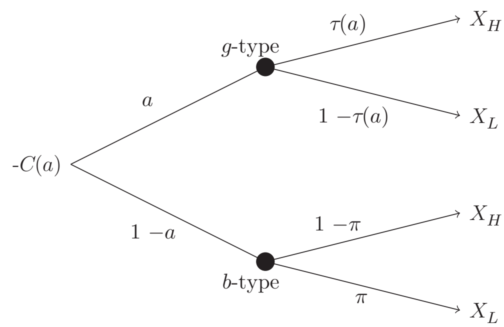

Figure 1. Project screening technology.

and that all projects have a positive net present value for the manager, who has a technology to privately screen project quality. By exerting effort $a \in [0,1]$ at nonpecuniary cost $C(a)$, where $C:[0,1] \to \mathbb{R}^+$, $C(0)=0$, $C'(0)=0$, and $C'(\cdot), C''(\cdot) > 0$ for $a \in (0,1]$, the manager can find a "good" project that pays X_H with probability $\tau(a)$ and X_L with probability $1-\tau(a)$, where $\tau:[0,1] \to [\pi,1]$. When a "good" project is found, it is financed by the manager since $\tau(a) \geq \pi$ for all $a \in [0,1]$. With probability $1-a$, however, screening is not successful and the manager finances an average project from the pool. Since the latter is a project the manager can finance without exerting any screening effort, it can be interpreted as the manager's outside option. The screening technology is depicted in Figure 1.

The screening technology assumed in this paper aims to capture banks' comparative advantage in loan origination. Evidence of banks being special lenders can be found in Fama (1985), and of banks having the ability to acquire private information to screen borrower quality in Mikkelson and Partch (1986), Slovin, Sushka, and Polonchek (1999), Plantin (2009), and Botsch and Vanasco (2016). A similar motivation applies when we think of active managers and large shareholders, who have an advantage in affecting the value of the firm as discussed in Shleifer and Vishny (1997).

After exerting effort a, the probability of obtaining cash flow X_H in $t=2$ is given by

$$\rho(a) \equiv \mathbb{P}_a(X = X_H) = a\tau(a) + (1-a)\pi. \tag{1}$$

I consider two cases: (1) *Endogenous Quality*: effort increases the expected cash flows of the "good" project, $\tau(0)=\pi$ and $\tau'(\cdot) > 0$, $\forall a \in [0,1]$, and (2) *Exogenous Quality*: the cash flows of the "good" project are independent of effort, $\tau(a) = \bar{\tau} > \pi$, $\forall a \in [0,1]$.

I assume that investors cannot observe the screening effort exerted by the manager nor the type of project that is financed. In practice, managers or originators exert screening effort by hiring better employees (e.g., loan officers), by devoting time to understand the pool of available projects, by designing optimal compensation packages, and in banking, by improving the technology used to verify the information content of loan applications. To the extent that the results of loan or asset screening are not deterministic, managers have private information about the screening technology and about the asset that is actually originated.

One of the contributions of this paper is the analysis of the endogenous-quality case. When quality is endogenous, effort affects both the likelihood of funding a good project and the profitability of this project. There are many settings in which this may be the case. By improving the search technology, a manager can increase the likelihood of finding a "better-than-average" project while at the same time increasing the quality of what is found. I show this formally in Appendix A, where I embed the manager with an information acquisition technology that can be used to screen a subset of projects. By investing in more precise information, the manager increases the likelihood of observing a good-signal project as well as the expected cash flows of this project, generating a structure such as that in Figure 1. I make the following assumption that will hold throughout the paper.

ASSUMPTION 1: *Functions $C(\cdot)$ and $\tau(\cdot)$ are such that:*

(i) $\exists\, \tilde{a} \in (0, 1)$ *such that* $\tilde{a} = \arg\max_{a\in[0,1]} \theta(\rho(a) - \pi)(X_H - X_L) - C(a)$.

(ii) $\frac{C'(a)}{\rho'(a)}$ *is increasing in* a *on* $(0, 1]$.

(iii) $\frac{C'(a)+C_R'(a)}{\rho'(a)}$ *is increasing in* a *on* $(0, 1]$*, where* $C_R(a) \equiv (\theta - 1)a\tau(a)\frac{C'(a)}{\rho'(a)}$.

The first condition ensures that the first-best level of effort is interior. The next two conditions state that the marginal cost of effort (including an indirect cost that will be described in Section II) increases faster than the marginal benefit of effort. This guarantees that the second-order conditions are satisfied both in the manager's problem and in the optimal mechanism.

Manager Types. The manager arrives at $t = 1$ with private information about her asset quality and her hidden action a. Let $z \in \{\mathbf{g}ood, \mathbf{b}ad\}$ denote the manager's type, where the g-type holds a good project and the b-type holds an average project.[7] Let $\mathbb{E}^z_a[\cdot]$ denote the expectations operator over the cash flows of the z-type manager that exerted effort a. When the expectation does not depend on effort a, I drop the subscript.[8]

Secondary Markets. At $t = 1$, the manager can sell a security backed by her asset cash flows. A security F is given by $t = 2$ payments contingent on cash flow realization: $F(X_L)$ and $F(X_H)$. I assume that the manager and investors have limited liability (LL), $0 \le F(X) \le X$, and I restrict attention to securities with payoffs that are weakly monotone in underlying cash flows (WM), $F(X_L) \le F(X_H)$ and $X_L - F(X_L) \le X_H - F(X_H)$.[9] I define the set of feasible securities as

[7] Effort a could be part of the manager's type. However, since in equilibrium it is unique and inferred by the market, it simplifies the problem to keep track of a and z separately.

[8] For any function $h : x \to h(x)$, we have $\mathbb{E}^g_a[h(X)] = \tau(a)h(X_H) + (1 - \tau(a))h(X_L)$ and $\mathbb{E}^b[h(X)] = \pi h(X_H) + (1 - \pi)h(X_L)$.

[9] These restrictions are taken from the security design literature, for example, Nachman and Noe (1994), DeMarzo and Duffie (1999), and Biais and Mariotti (2005). For a rich discussion on the motivation to impose these restrictions and their implications, see Innes (1990).

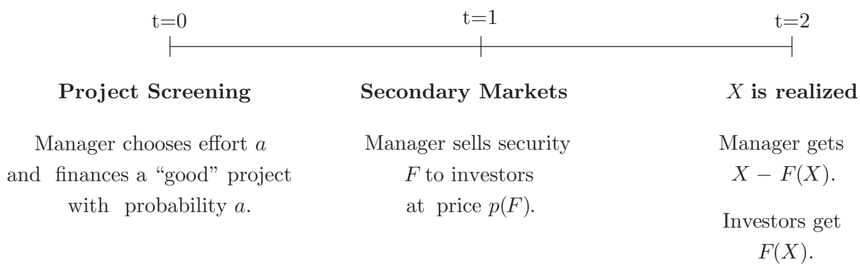

Figure 2. Timeline of the model.

$\Delta \equiv \{F : \text{(LL) and (WM) hold}\}$. The manager offers to sell security $F \in \Delta$ to investors, and thus I solve a signaling game in secondary markets. I apply the standard equilibrium selection by using the D1-Refinement criterion.[10]

The timing of the game is presented in Figure 2.

B. The Manager's Problem and Equilibrium Definitions

The manager solves two problems. In $t = 1$, a z-type manager chooses what cash flows to sell to investors. In $t = 0$, given secondary market strategies, the manager chooses how much effort to exert to screen projects. The model therefore features a problem of incentives at the asset screening stage, followed by a problem of asymmetric information at the asset cash flows sale stage. I begin by describing the manager's problem in $t = 1$, followed by the one in $t = 0$.

The strategy of a z-type manager in secondary markets is given by the security that she sells to investors. Let $p : \Delta \to \mathbb{R}^+$ be the mapping from securities to the price paid by investors. The problem of a z-type manager in $t = 1$ given effort a chosen in $t = 0$ is

$$\max_{F \in \Delta} \ \theta p(F) + \mathbb{E}_a^z[X - F(X)]. \tag{2}$$

Rather than defining the strategies of investors, I model the buyer side of the market as a price function for all feasible securities. Since the market is competitive, this pricing function needs to ensure that investors make zero profits in expectation.

Investors form beliefs about the manager's screening action, denoted by a^e, and about the manager's z-type, denoted by $\mu : \Delta \to [0, 1]$, where $\mu(F)$ is the probability of a manager being a g-type if she chooses to sell security F. As a result, the market valuation for security $F \in \Delta$ is denoted by $\mathbb{E}_{a^e}^{\mu}[F(X)]$, where

$$\mathbb{E}_{a^e}^{\mu}[F(X)] = \mu(F)\mathbb{E}_{a^e}^{g}[F(X)] + (1 - \mu(F))\mathbb{E}^{b}[F(X)]. \tag{3}$$

[10] See Cho and Kreps (1987) and Cho and Sobel (1990).

In what follows, I define equilibria in secondary markets.

DEFINITION 1: *Given any level of effort a and market beliefs* a^e*, an equilibrium in secondary markets is given by a pricing function* $p : \Delta \to \mathbb{R}^+$*, a manager z-type strategy* $\sigma(z) = \{F_z(X_L), F_z(X_H)\}$ *for* $z \in \{g, b\}$*, and belief function* $\mu : \Delta \to [0, 1]$ *satisfying the following conditions:*

1. *Manager's Optimality: Given* $p(\cdot)$*,* $\sigma(z)$ *is the solution to* (2) *for* $z \in \{g, b\}$*.*
2. *Belief Consistency:* $\mu(\cdot)$ *is derived from* $\sigma(\cdot)$ *using Bayes's rule when it applies.*
3. *Zero Profit Condition:* $p(F) = \mathbb{E}^{\mu}_{a^e}[F(X)]$*.*

A secondary market equilibrium outcome is a set of prices and securities per manager type, denoted by $\Phi(a, a^e) \equiv \{p_z, F_z\}_{z\in\{g,b\}}$. Prices and securities are a function of both $\{a, a^e\}$, but I drop this indexing to save on notation. Given $\Phi(\cdot, \cdot)$, the value to the manager in $t = 0$ is

$$V_0(a, a^e) = a(\theta p_g + \mathbb{E}^g_a t[X - F_g(X)]) + (1 - a)(\theta p_b + \mathbb{E}^b[X - F_b(X)]) - C(a). \quad (4)$$

I focus on pure-strategy equilibria on the choice of effort a, and I assume that market beliefs about this action are thus degenerate at some value $a^e \in [0, 1]$. I next define the equilibrium of the full game.

DEFINITION 2: *An equilibrium is given by* $\{a^e, a^*, p^*_g, p^*_b, F^*_g, F^*_b\} \in [0, 1]^2 \times \mathbb{R}^2_+ \times \Delta^2$ *satisfying the following conditions:*

1. *Manager's Optimality in* $t = 0$*: Given* a^e *and* $\Phi(\cdot, a^e)$*,* $a^*(a^e) = \arg\max_{a\in[0,1]} V_0(a, a^e)$*.*
2. *Secondary Market Equilibrium:* $\{p^*_z, F^*_z\}_{z\in\{g,b\}} = \Phi(a^*, a^*)$*.*
3. *Belief Consistency:* $a^e = a^* = a^*(a^*)$*.*

C. The First-Best

I characterize the first-best of this economy as a useful benchmark for the remainder of the paper. The results are presented in the following proposition.

PROPOSITION 1: *In the first-best, the manager sells a full claim to her cash flows to investors in* $t = 1$*,* $F^*_{FB}(X) = X$*, and exerts effort* $a^*_{FB} > 0$ *in* $t = 0$ *given by:*

$$\theta \rho'(a^*_{FB})(X_H - X_L) - C'(a^*_{FB}) = 0. \quad (5)$$

First-best allocations are obtained in the equilibrium with observable effort and z-type.

In the first-best, all $t = 2$ cash flows are sold to investors in $t = 1$, since the manager values funds in $t = 1$ more, independent of her z-type and effort choice. Equivalently, in the equilibrium with full information, gains from trade are maximized when all cash flows are sold, since there are no costs associated with selling cash flows. In the first-best, the manager's effort choice equalizes the social marginal benefit of effort with its social marginal cost. This is also true in the full information equilibrium, since the manager is fully compensated for her effort choice in secondary markets. Interestingly, contractability on the effort choice is not important, as observability of a and z is enough to implement socially efficient allocations. Finally, from (5) we have that a^*_{FB} is increasing in θ.

In what follows, I explore how these allocations may be distorted in the presence of information asymmetries.

II. Optimal Mechanism

In this section I characterize the optimal mechanism that maximizes ex ante efficiency. The results from this benchmark will highlight the inefficiencies (relative to first-best) that arise when screening effort and the manager's z-type are not observed by investors. In Appendix F, I analyze alternative information structures, such as the case of observable a and unobservable z, and vice versa.

I focus on direct revelation mechanisms that stipulate a $t = 1$ transfer $t_{\hat{z}}$ contingent on reported type $\hat{z} \in \{g, b\}$, and $t = 2$ transfers $T_{\hat{z}}(x)$ contingent on both reported z-type, $\hat{z}$, and cash flow realization, $x \in \{X_L, X_H\}$.[11] The transfers are financed by investors, who receive full ownership of $t = 2$ cash flows. The mechanism therefore needs to satisfy investors' participation constraint. Finally, the level of effort implemented by the mechanism has to be incentive compatible (IC). This is stated formally in the following definition.

DEFINITION 3: *The optimal mechanism is given by an implementable effort level and transfers* $\{a, t_z, T_z(\cdot)\}_{\{z \in g, b\}}$ *that maximize the value for the manager in* $t = 0$

$$\max_{\{a, t_z, T_z(\cdot)\}} a[\theta t_g + E^g_a[T_g(X)]] + (1 - a)[\theta t_b + E^b[T_b(X)]] - C(a), \tag{6}$$

subject to:

1. *Incentive Compatibility for Type Revelation:*

$$\theta t_g + E^b[T_g(X)] \leq \theta t_b + E^b[T_b(X)] \tag{7}$$

$$\theta t_b + E^g_a[T_b(X)] \leq \theta t_g + E^g_a[T_g(X)]. \tag{8}$$

2. *Investors' Participation Constraint:*

$$a[-t_g + E^g_a[X - T_g(X)]] + (1 - a)[-t_b + E^b[X - T_b(X)]] \geq 0. \tag{9}$$

3. *Incentive Compatibility for Effort Choice:*

$$\begin{aligned} a = \arg\max_{\hat{a}} \ & \hat{a} \max\left\{\theta t_g + E^g_{\hat{a}}[T_g(X)], \theta t_b + E^g_{\hat{a}}[T_b(X)]\right\} \\ & + (1 - \hat{a})\left(\theta t_b + E^b[T_b(X)]\right) - C(\hat{a}). \end{aligned} \tag{10}$$

[11] By the Revelation Principle, we know that, for any Bayesian–Nash equilibrium, there exists a direct mechanism that is payoff-equivalent and where truthful revelation is an equilibrium.

4. *Feasibility:* $T_z \in \Delta$ *for* $z \in \{g, b\}$.

This problem is similar to the one presented in Biais and Mariotti (2005). They study optimal mechanism design in the presence of adverse selection, where an issuer with private information about asset quality has to issue a security to uninformed competitive liquidity providers. In contrast to their paper, in this setup the quality of underlying assets and of the private information held by the manager depend on the screening effort, which is her hidden action. As a result, the mechanism designer internalizes the effect that different transfers have on incentives.

Transfers and Retention. The $t = 2$ transfers should be interpreted as the retention of cash flows in the optimal mechanism, where $F_z(X) = X - T_z(X)$ is the security sold by the z-type manager. Consistent with this interpretation, t_z is the price received for security F_z, and T_z is the retention of cash flows in the optimal mechanism.

Global Deviations. Constraint (10) controls for the possibility of the manager choosing to deviate on her effort in $t = 0$ and misreporting her z-type in $t = 1$. This deviation can arise only when quality is endogenous, since the g-type incentive compatibility constraint holds at the implemented effort level, but need not hold if the manager chooses to deviate. To address this, I proceed as follows. I replace the incentive compatibility constraint for effort (10) with the first-order condition for effort choice, obtained when the incentive compatibility for type revelation of the g-type (8) holds:

$$\underbrace{\theta(t_g - t_b) + E_a^g[T_g(X)] - E^b[T_b(X)]}_{\text{Difference in } t\text{=1 Payoff between } g\text{- and } b\text{-types}} + \underbrace{a\tau'(a)\left[T_g(X_H) - T_g(X_L)\right]}_{\text{Marginal Change in Quality of Retention}}$$
$$= \underbrace{C'(a).}_{\text{Marginal Cost}} \tag{11}$$

Later, I verify that the allocations obtained under the first-order approach satisfy global incentive compatibility.

The following lemma presents the first important result: only retention of the g-type manager is desired in the optimal mechanism.

LEMMA 1: *Under the optimal mechanism, the bad-type manager does not retain any cash flows,* $T_b(X_L) = T_b(X_H) = 0$*, while the good-type manager retains a junior claim to her cash flows,* $T_g(X) = \max\{0, X - d\}$ *for* $d \in [X_L, X_H]$ *so* $T_g(X_L) = 0$ *and* $T_g(X_H) = X_H - d$.

I refer to d as the debt level since $X - T_g(X) = \min\{d, X\}$ can be interpreted as the debt security sold by the g-type manager in secondary markets. Using the results from Lemma 1, the following lemma characterizes the $t = 1$ transfers, which are pinned down by the binding incentive compatibility constraint of the b-type manager and the binding participation constraint of investors.

LEMMA 2: *Under the optimal mechanism, for given effort and debt levels* $\{a, d\}$, *the* $t = 1$ *transfers are given by*

$$t_b = aE_a^g[\min\{d, X\}] + (1 - a)E^b[X] + a\frac{\pi}{\theta}(X_H - d) \tag{12}$$

$$t_g = aE_a^g[\min\{d, X\}] + (1 - a)E^b[X] - (1 - a)\frac{\pi}{\theta}(X_H - d). \tag{13}$$

Retention of cash flows is essential to implement positive effort. If retention is zero for both types, $T_z(X) = 0$, then $t = 1$ transfers, t_z, have to be equal for both types by the incentive compatibility constraints. Thus, from (11), the manager has no incentives to exert screening effort. Even though retention is necessary, Lemma 1 shows that imposing the same cash flow retention on all manager types is inefficient. Equal retention levels reduce gains from trade without necessarily improving incentives. The manager has two motives to exert effort: (i) to increase the probability of financing a good project, and (ii) to improve the quality of the cash flows that she expects to retain. From (11), these motives are strengthened by increasing the expected retention of the g-type relative to that of the b-type, $E_a^g[T^g(X)] - E^b[T^b(X)]$, and by increasing the g-type differential payoff across states: $T_g(X_H) - T_g(X_L)$. As a result, the b-type retains zero and the g-type retains a junior claim to her cash flows.

Finally, type-contingent transfers in $t = 1$ are chosen to ensure a binding participation constraint of investors (all surplus is transferred to the manager), and a binding incentive compatibility for the b-type (all surplus is transferred to the g-type subject to the b-type not mimicking). As a result, the mechanism can separate types for any positive retention level. The results from Lemmas 1 and 2 fully characterize the optimal mechanism transfers as a function of the implementable effort level $a \in [0, 1]$ and debt level $d \in [X_L, X_H]$. The following corollary incorporates these results and states how effort and debt levels are determined.

COROLLARY 1: *In the optimal mechanism, effort and debt level* $\{a^*, d^*\}$ *solve:*

$$\max_{a\in[0,1],d\in[X_L,X_H]} \theta E_a[X] - a\tau(a)(\theta - 1)(X_H - d) - C(a) \tag{14}$$

$$\textit{subject to:} \qquad \rho'(a)(X_H - d) = C'(a). \tag{15}$$

Screening effort is always below first-best: $a^* < a^*_{FB}$.

Corollary 1 shows that effort is chosen to maximize the manager's $t = 0$ value, which is lower than in the first-best when there is retention: $X_H - d$. In addition, constraint (15) shows that, to implement a given level of effort, retention is required. It is useful to define the indirect cost of effort given by the retention required to implement it

$$C_R(a) \equiv a\tau(a)(\theta - 1)\frac{C'(a)}{\rho'(a)}, \tag{16}$$

where $a\tau(a)$ is the probability of retention, $\theta - 1$ is the marginal cost of retention, and $\frac{C'(a)}{\rho'(a)}$ is the retention required to implement effort a as stated in (15). Finally, let $\bar{a} \in (0, 1]$ be the maximum effort level that can be implemented under the optimal mechanism, given by $\rho'(\bar{a})(X_H - X_L) = C'(\bar{a})$. By comparison with (5), it follows that the level of effort under the optimal mechanism is always below the first-best, $\bar{a} < a^*_{FB}$.

The following condition is necessary and sufficient for positive effort to be implemented under the optimal mechanism. The condition states that there exists a positive effort level that gives the manager a higher $t = 0$ payoff than exerting zero effort.

CONDITION 1: *There exists $\tilde{a} \in (0, \bar{a}]$ such that*

$$\theta(\rho(\tilde{a}) - \pi)(X_H - X_L) - C(\tilde{a}) - C_R(\tilde{a}) > 0. \tag{17}$$

The following proposition combines the results from Lemmas 1 and 2 and from Corollary 1 and concludes the characterization of the optimal mechanism.

PROPOSITION 2: *If Condition 1 holds, the optimal mechanism's effort and debt levels $\{a^*, d^*\}$ are given by $a^* > 0$ and $d^* < X_H$ such that*

$$(a^*, d^*) = \begin{cases} \left(a_{int}\ ,\ X_H - \frac{C'(a_{int})}{\rho'(a_{int})}\right) & if \quad a_{int} \leq \bar{a} \\ (\bar{a}\ ,\ X_L) & otherwise, \end{cases} \tag{18}$$

where a_{int} is the interior solution to (14), and is given by

$$\theta\rho'(a_{int})(X_H - X_L) - C'(a_{int}) - C'_R(a_{int}) = 0. \tag{19}$$

If Condition 1 does not hold, the optimal mechanism implements $a^* = 0$ and $d^* = X_H$. Transfers $\{t^*_b, t^*_g, T^*_b(\cdot), T^*_g(\cdot)\}$ are given by Lemmas 1 and 2.

The optimal mechanism chooses effort such that the social marginal benefit of exerting effort equals its social marginal costs. When comparing the level of effort under the optimal mechanism given by (19) with the first-best level given by (5), we see that the indirect cost generates a wedge between the two effort levels. This is because costly retention is required to implement effort under the optimal mechanism.[12] It is important to highlight that this mechanism uses the manager's private information to implement differential retention levels. Indeed, the mechanism achieves higher ex ante welfare when the manager has private information about her asset quality than when she does not (see Appendix F).

The optimal mechanism can implement first-best allocations if effort a or the manager's z-type are verifiable. The results are intuitive and are presented in Appendix F. When z-types are verifiable, the optimal mechanism can use

[12] These results apply for both the endogenous- and the exogenous-quality case, since allocations are characterized as a function of the marginal benefits and costs of effort. For a description of the nature of these marginal benefits and costs and how they differ in each case, see Appendix B.

type contingent $t = 1$ transfers to implement first-best effort levels, and costly retention is no longer necessary. When screening effort is verifiable, $t = 1$ transfers are constant across types and ensure zero retention and first-best effort level is implemented by paying zero after observed deviations. Thus, it is the combination of unobservable effort and z-types that drives the wedge between first- and second-best allocations. This highlights the importance of analyzing moral hazard on screening effort and asymmetric information on asset types simultaneously.

Comparative Statics. The following proposition establishes how changes in gains from trade or liquidity needs, in the cost of exerting effort, and in the projects' cash flows affect the levels of effort and debt under the optimal mechanism.

PROPOSITION 3: *Suppose that effort level a^* and debt level d^* implemented in the optimal mechanism are interior. Then, we have the following comparative statics:*

1. *Gains from Trade: a^* is decreasing and d^* is increasing in θ.*
2. *Cost of Effort: suppose that $C(\cdot) = \chi h(\cdot)$ for some $\chi > 0$. Then, a^* is decreasing in χ and d^* can be either increasing or decreasing in χ.*
3. *Differential Quality: a^* and retention $X_H - d^*$ are increasing in $X_H - X_L$.*

First, when gains from trade increase there are two opposing effects: on the one hand, retention becomes more costly; on the other hand, the return to effort increases. I show that the first force always dominates, and that the optimal mechanism always reduces retention at the expense of lower effort. Note that the incentive compatibility constraint for effort is independent of θ, since the manager does not internalize the increase in the return to effort due to a larger θ. The manager does, however, internalize the increase in the indirect cost of implementing effort. This is in sharp contrast to the first-best, where effort levels are increasing in the gains from trade since the return on effort and the cost of retention are both internalized by the manager. As a result, gains from trade increase the wedge between first- and second-best effort levels. Second, when the marginal cost of effort increases, effort decreases. The effect on debt levels, however, is ambiguous. On the one hand, retention should decrease since lower effort needs to be implemented. On the other hand, effort is costlier to implement since higher retention is needed to obtain a given effort level. As a result, depending on which force dominates, debt levels can decrease or increase. Finally, an increase in the cash flow differential between high and low states increases the return to effort, increasing the optimal effort choice and thus the required retention.

III. Market Equilibria

In this section I characterize the market equilibrium allocations of Definitions 1 and 2. In contrast to the optimal mechanism, the manager chooses which cash flows to sell to maximize her $t = 1$ value, at which point effort a

and investors' beliefs a^e are given. At $t = 0$, the manager only chooses how much effort to exert, given the optimal strategies in secondary markets for each z-type. It is implicitly assumed that the manager and investors cannot commit to choices made in $t = 0$. Thus, the results of this section describe the market equilibrium allocations that arise when there is a lack of commitment. In Section IV, I discuss the role of commitment of the manager and investors.

I solve the model by backwards induction. First, for a given pair $\{a, a^e\}$, I solve the z-type manager's problem in $t = 1$. Given secondary market optimal strategies, I then solve the manager's problem in $t = 0$.

PROPOSITION 4: *Let $\{a, a^e\}$ be given. There exists a unique equilibrium in secondary markets where the b-type manager sells $F_b(X) = X$ and the g-type manager sells $F_g(X) = \min\{d, X_L\}$, where*

$$d(a, a^e) = \begin{cases} X_L + \frac{(\theta-1)\pi}{\theta\tau(a^e)-\pi}(X_H - X_L) & \theta\tau(a^e) \geq \tau(a) \\ X_L & \theta\tau(a^e) < \tau(a). \end{cases} \tag{20}$$

When $a = a^e$, the least costly separating equilibrium (LCSE) is the only equilibrium that survives the D1-Refinements. The b-type manager sells a full claim to her cash flows and receives her full information payoff, while the g-type manager sells a debt-like security, where debt levels are determined by the binding incentive compatibility of the b-type manager.[13] Note that Definition 1 allows for the level of effort a to differ from market beliefs a^e.

The results presented in Proposition 4 can be summarized as follows. When the manager has chosen effort $a \leq a^e$ or when gains from trade are sufficiently large, the g-type manager sells the same security she would have sold if $a = a^e$ (i.e., LCSE strategies). This is because when $\theta\tau(a^e) > \tau(a)$, the g-type manager prefers to sell, even when investors undervalue her cash flows. However, when the manager has chosen to implement a higher effort level $a > a^e$ and gains from trade are not large enough, the manager sells a risk-free claim in secondary markets. That is, the manager deviates from the LCSE strategies when $\theta\tau(a^e) < \tau(a)$, since cash flows are so undervalued by investors that the manager prefers not to sell. Since the value of a risk-free claim is independent of market beliefs, investors do not need to form beliefs about the manager's type to price this security. Note that this can only happen off the equilibrium path (since $a > a^e$), but it will be important in characterizing the choice of effort and the existence of equilibrium of the full game. The following proposition characterizes debt and effort levels in any equilibrium.

PROPOSITION 5: *In any equilibrium, effort and debt levels $\{a, d\}$ must satisfy the following two conditions:*

$$\rho'(a)(X_H - d) - C'(a) = 0 \tag{21}$$

[13] This result is analogous to DeMarzo (2005).

$$d = X_L + \frac{(\theta - 1)\pi}{\theta\tau(a) - \pi}(X_H - X_L). \tag{22}$$

When quality is endogenous, there are at least two solutions to (21)–(22): one with $a > 0$ and $d < X_H$ and another with $a = 0$ and $d = X_H$. When quality is exogenous, there is a unique solution to (21)–(22) in which $a > 0$.

From now on, I denote the solution to (21)–(22) that yields positive effort by $\{a_M^*, d_M^*\}$, and the solution with zero effort by $\{0, X_H\}$. I provide the conditions under which there are at most two solutions to (21)–(22) in Appendix E, and I restrict my attention to this case henceforward. The results in Proposition 5 are obtained by solving the manager's problem in $t = 0$, where effort is chosen given secondary market outcomes $\Phi(\cdot, a^e)$:

$$a^*(a^e) = \arg\max_a \;\; a(\theta p_g + \tau(a)(X_H - d(a, a^e))) + (1 - a)\theta p_b - C(a), \tag{23}$$

where $p_g = \tau(a^e)(d(a, a^e) - X_L) + X_L$, $p_b = \pi(X_H - X_L) + X_L$, $d(a, a^e)$ given by (20), and where condition $a^* = a^*(a^*) = a^e$ is imposed.

Proposition 5 presents the two main results of this section. First, effort is chosen such that the private marginal benefit of effort to the manager, which is directly proportional to the amount retained, equals the marginal cost to the manager (equation (21)). Second, the amount retained in secondary markets is fully pinned down by the amount that is needed for the g-type manager to signal her quality to investors (equation (22)). Therefore, the presence of asymmetric information is what sustains the equilibrium with positive effort and asset screening.[14] In Appendix F, I show that if the manager did not have private information about her asset (i.e., there were no manager types), the unique equilibrium features zero retention and no screening effort. This is because, with no private information about project quality, the manager is able to sell a full claim to her cash flows in $t = 1$. Since cash flows are priced with investors' beliefs about the screening effort instead of the manager's actual choice, the manager has no incentives to screen projects in $t = 0$.

A detailed comparison of the market equilibrium allocations with those of the optimal mechanism is provided in Section IV.

Comparative Statics. The following proposition describes how effort and debt levels vary with gains from trade, cost of effort, and differential quality of the cash flows. The comparative statics focus on the equilibrium candidate with positive effort.

PROPOSITION 6: *The equilibrium effort and debt levels $\{a_M^*, d_M^*\}$ with $a_M^* > 0$ have the following comparative statics:*

1. *Gains from Trade: a_M^* is decreasing and d_M^* is increasing in θ.*

[14] In Appendix F, I show that, in the model with no private information (i.e., the manager exerts hidden effort but does not observe her z-type), the unique equilibrium features no screening effort and no retention.

2. *Cost of Effort: Suppose that $C(a) = \chi h(a)$ for some $\chi > 0$. Then a_M^* is decreasing in χ, while d_M^* is increasing in χ when quality is endogenous and constant when quality is exogenous.*
3. *Differential Quality: a_M^* and $X_H - d_M^*$ are increasing in $X_H - X_L$.*

First, effort levels are decreasing in gains from trade, while debt levels are increasing, as in the optimal mechanism, since in both cases higher gains from trade increase the cost of retention. In contrast to the optimal mechanism, however, gains from trade do not affect the return to effort, since the manager does not internalize the effect of her effort on market prices. Thus, as gains from trade increase, equilibrium allocations feature lower retention levels, which lower ex ante incentives to screen projects. This behavior is in line with the observed trend of securitized assets, such as nonagency MBS, that featured a boom in the years leading up to the crisis due to increasing demand for securitized products, that was accompanied by a decrease in the quality of the underlying loans (Mian and Sufi (2009)).

Second, higher costs of effort reduce screening incentives and, as a result, retention in secondary markets. The decline in effort is a direct response to the increase in its marginal cost. The response of debt levels, however, occurs only when quality is endogenous. In this case, lower effort reduces information asymmetries between the manager and investors, allowing the former to sell more cash flows. When quality is exogenous debt levels do not depend on the level of effort and thus retention remains constant. Finally, when differential cash flow quality increases, the return to effort increases for a given retention level. In addition, both the increase in effort and in quality differential exacerbates the asymmetric information problem, increasing the level of cash flow retention needed for the g-type to separate.

A. Existence and Multiplicity of Market Equilibria

To establish the existence of an equilibrium with no commitment, we need to rule out double-deviations. First, note that deviations to lower effort levels are ruled out since by Proposition 4 they imply issuing the LCSE debt level in secondary markets and thus are consistent with on-equilibrium-path effort choices in $t = 0$. Thus, we need to rule out deviations to higher effort levels accompanied by the issuance of a risk-free security in secondary markets. Let $\tilde{a} \in [0, 1]$ denote the optimal choice of effort when the manager expects to issue a risk-free security in secondary markets if it originates a "good" project, that is, $F_g(X) = X_L$. This best-deviation on effort $\tilde{a}$ is given by

$$\tilde{a} = \arg\max_{a \in [0,1]} (a\tau(a) + (1 - a)\theta\pi)(X_H - X_L) + \theta X_L - C(a). \tag{24}$$

The following proposition establishes that when asset quality is endogenous there can be one or two equilibria or an equilibrium can even fail to exist, while with exogenous quality an equilibrium always exists and is unique.

PROPOSITION 7: *[Endogenous Quality] Let $\{a_M^*, d_M^*\}$ and $\{0, X_H\}$ be the two equilibrium candidates in Proposition 5. The following conditions characterize the existence of equilibria:*

1. *If $\theta\pi \geq \tau(\tilde{a})$, then both $\{a_M^*, d_M^*\}$ and $\{0, X_H\}$ are equilibria.*
2. *If $\theta\pi < \tau(\tilde{a})$ and $\theta\tau(a_M^*) \geq \tau(\tilde{a})$, then $\{a_M^*, d_M^*\}$ is the unique equilibrium.*
3. *If $\theta\tau(a_M^*) < \tau(\tilde{a})$, then $\{a_M^*, d_M^*\}$ is the unique equilibrium if $V_0(a_M^*, d_M^*) \geq V_0(\tilde{a}, X_L)$; otherwise, an equilibrium does not exist.*

[Exogenous Quality] The candidate $\{a_M^*, d_M^*\}$ in Proposition 5 is the unique equilibrium.

Proposition 7 gives rise to the following corollary that characterizes the equilibrium set for the endogenous quality case as a function of the gains from trade, θ.

COROLLARY 2: *When quality is endogenous, there exists $\{\underline{\theta}, \bar{\theta}\}$, where $1 < \underline{\theta} < \bar{\theta}$, such that for $\theta \geq \bar{\theta}$ there are multiple equilibria, and for $\theta < \underline{\theta}$ an equilibrium may fail to exist. The equilibrium is unique when $\theta \in [\underline{\theta}, \bar{\theta})$.*

The results in Corollary 2 follow directly from Proposition 7, since $\tau(\cdot)$ increases in the level of effort, $\tilde{a}$ decreases as gains from trade, θ, increase by (24), and a_M^* is increasing in θ by Proposition 6.

When quality is endogenous and gains from trade are large enough so that $\theta\pi \geq \tau(\tilde{a})$, one of two equilibria can arise: one with positive screening effort and retention of cash flows, and one with zero screening effort and no retention of cash flows. Market beliefs are self-fulfilling. If the market believes the manager has exerted no effort, since $\theta\pi - \tau(\tilde{a}) > 0$, the manager's best response is to sell all cash flows, which is consistent with exerting no effort. In contrast, if the market believes that the manager has exerted effort $a_M^* > 0$, since $\theta\tau(a_M^*) - \tau(\tilde{a}) > 0$, the g-type manager's best response is to issue debt d_M^* and retain the remaining cash flows, which is consistent with exerting effort a_M^* ex ante.

When quality is endogenous but gains from trade are not large enough so that $\theta\tau(a_M^*) < \tau(\tilde{a})$, an equilibrium with endogenous quality may fail to exist. This highlights the instability of a market where both the quality of asset screening and the level of private information held by the asset originator are connected through a hidden effort choice. Since retention is valuable ex ante to generate incentives, when the manager lacks commitment, by exerting more effort than what the market expects, the manager creates a commitment device to not-sell in secondary markets. As a result, she exerts effort ex ante and this may provide higher ex ante value.

Finally, when quality is exogenous an equilibrium always exists and is unique. The reason is that, in this case, the level of information asymmetries in secondary markets is independent of effort, and so is the retention of cash flows.

IV. Ex Ante Efficiency and the Role of Policy

The key difference between the optimal mechanism of Section II and the market equilibrium allocations of Section III is the ability to commit at $t = 0$ to choices that are made in $t = 1$. With lack of commitment, the manager chooses how much to retain when effort is a sunk action, and transfers in secondary markets across manager types are zero since the equilibrium is separating. The lack of commitment introduces externalities that can be addressed with policy. However, if the manager and investors can commit in $t = 0$ to $t = 1$ choices, second-best allocations obtain and there is no need for policy intervention.

The contracts between the manager and investors that need to be written to implement second-best are not necessarily renegotiation proof. After effort is exerted, there are gains from trade to be exploited. Therefore, the lack of commitment may be a prevalent friction in many markets. On the other hand, there exist market mechanisms that may address the problem of commitment to cash flow retention. One example comprises "master repurchase agreements," where originators commit to buy back certain nonperforming loans from the buyer at later dates. However, Ashcraft, Gooriah, and Kermani (2014) find that, in the conduit CMBS market, the level of cash flow retention cannot be predicted by prices at origination, suggesting the absence of predetermined contracts. Overall, the presence or lack of commitment in a given market remains an empirical question.

In what follows, I study how informational frictions drive a wedge between the first- and second-best allocations, and how market equilibrium allocations differ from the second-best allocations when markets lack commitment. The following comparison is helpful to understand the different constraints present in each problem. Under the full-information benchmark (first-best), under the optimal mechanism (second-best), and for market equilibrium allocations, the value for the manager in $t = 0$ can be expressed as[15]

$$V_0(a, d) = \theta E_a[X] - C(a) - (\theta - 1)a\tau(a)(X_H - d). \tag{25}$$

The first-best maximizes (25) without any constraints. The second-best maximizes (25) subject to the incentive compatibility constraint for effort,

$$\rho'(a)(X_H - d) = C'(a). \tag{26}$$

Finally, the market allocations are pinned down by the incentive compatibility constraint for effort (26) and the additional LCSE constraint:

$$d = X_L + \frac{(\theta - 1)\pi}{\theta\tau(a^e) - \pi}(X_H - X_L), \tag{27}$$

together with the equilibrium condition $a^e = a^*$. After some algebra, we have that effort levels in each case are determined by

$$\theta\rho'(a^*_{FB})(X_H - X_L) = C'(a^*_{FB}) \tag{28}$$

[15] In each case, the expression is obtained by plugging in the corresponding prices and transfers, where $d = X_H$ in the full-information benchmark.

$$\theta\rho'(a^*)\mathcal{D}_{OM}(a^*)(X_H - X_L) = C'(a^*) \tag{29}$$

$$\theta\rho'(a_M^*)\mathcal{D}_{Mk}(a_M^*)(X_H - X_L) = C'(a_M^*), \tag{30}$$

where $\mathcal{D}_{OM}(a)$ and $\mathcal{D}_{Mk}(a)$ capture the distortions relative to the first-best under the optimal mechanism and the market equilibrium allocations, respectively, and are given by

$$\mathcal{D}_{OM}(a) \equiv \frac{\tau(a) - \pi + ee(a)}{\theta\tau(a) - \pi + ee(a) + re(a)} \tag{31}$$

$$\mathcal{D}_{Mk}(a) \equiv \frac{\tau(a) - \pi}{\theta\tau(a) - \pi}, \tag{32}$$

where $ee(a) \equiv a\tau'(a)$, $re(a) \equiv (\theta - 1)\tau(a)(\eta_\tau(a) + \eta_c(a) - \eta_{\rho'}(a))$ for $a \in (0, 1]$, $\eta_\tau(a) \equiv \frac{a\tau'(a)}{\tau(a)}$, $\eta_{\rho'}(a) \equiv \frac{a\rho''(a)}{\rho'(a)}$, and $\eta_c(a) \equiv \frac{aC''(a)}{C'(a)}$ are the elasticities of $\tau(\cdot)$, $\rho'(\cdot)$, and $C(\cdot)$ with respect to effort a, and where $\eta_c(a) - \eta_{\rho'}(a) \geq 0$ for all $a \in (0, 1]$ by Assumption 1.

Second-best allocations differ from the first-best due to the indirect cost, since cash flow retention is needed to implement screening effort. In the market equilibrium, lack of commitment introduces two externalities. First, the manager no longer internalizes the effect of her screening effort on the quality of the cash flows that are sold to investors. This is because prices for cash flows that are sold are a function of market beliefs, a^e. I refer to this as an *effort externality*, which is captured by the term $ee(a)$ and is only present when quality is endogenous. Second, the manager does not internalize how the choice of retention affects her ex ante value. I refer to this as a *retention externality*, which is measured by $re(a)$ and is present whenever there is adverse selection in secondary markets. While the effort externality leads the manager to exert too little effort relative to the second-best, the retention externality leads the manager to exert too much effort relative to the second best, since retention of cash flow is inefficiently high. Therefore, depending on which force dominates, the market may feature more or less retention and in turn effort relative to the second-best.

REMARK 1: *When quality is exogenous, the market features excessive effort relative to second-best: $a_M^* > a^*$ and $d_M^* < d^*$.*

When quality is exogenous, there is no effort externality. The manager fully internalizes the direct return to her effort choice. The retention externality, however, is present since good-type managers retain ex post due to adverse selection. Thus, $ee(a) = 0$ while $re(a) = (\theta - 1)\bar{\tau}\eta_c(a) > 0$, $\forall a \in (0, 1)$. The optimal mechanism therefore features lower cash flow retention and screening effort than the market equilibrium.

REMARK 2: *When quality is endogenous, the market features lower screening effort relative to the second-best allocations when it is in the bad equilibrium*

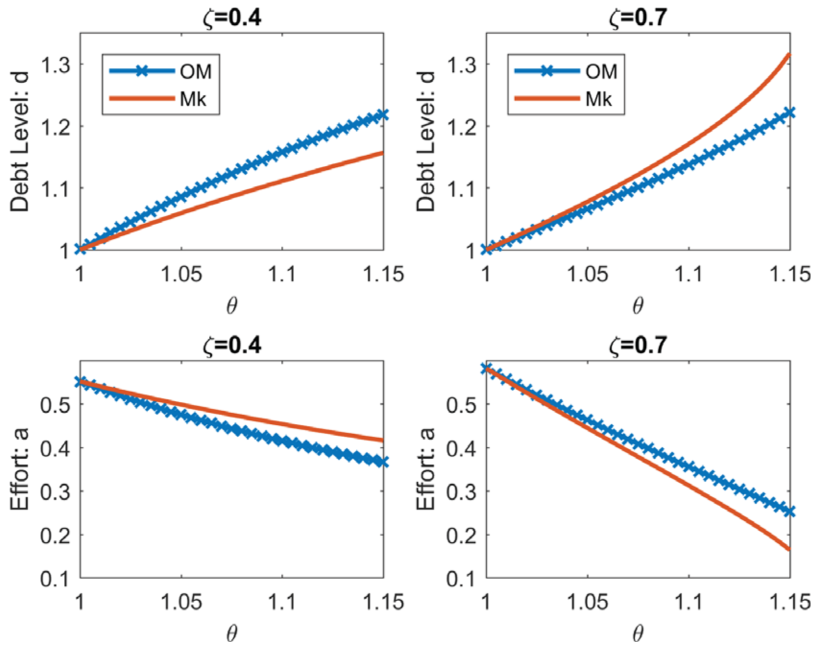

Figure 3. Optimal mechanism (OM) and market equilibrium allocations (Mk). This figure plots the optimal cash flow retention given by debt level d (two top panels) and the optimal choice of effort a (two bottom panels) under the optimal mechanism and the market allocations for different gains from trade θ. Results are for the endogenous quality case with functional form $\tau(a) = \pi + a^{\zeta}(1-\pi)$. The two panels on the left are computed with $\zeta = 0.4$ and the two on the right with $\zeta = 0.7$. In all panels $C(a) = \frac{a^2}{2}$, $\pi = 0.5$, $X_H = 2$, and $X_L = 1$. (Color figure can be viewed at wileyonlinelibrary.com)

with no screening and no retention, $\{0, X_H\}$, *or when it is in the good equilibrium* $\{a_M^*, d_M^*\}$ *and the effort externality is stronger than the retention externality.*

It is clear that the optimal mechanism implements more screening effort and more retention than the market when the latter is in the bad equilibrium, $\{0, X_H\}$. Otherwise, the optimal mechanism implements higher cash flow retention to improve screening effort when the effort externality dominates, that is, when $ee(a_M^*)$ is sufficiently larger relative to $re(a_M^*)$. By inspection of (31), this will depend on the curvature of the $\tau(\cdot)$ function, that is, how sensitive the asset quality is to screening effort. In particular, the effort externality will tend to dominate when $\tau'(\cdot)$ is large. This is illustrated in Figure 3, where debt and effort levels are plotted for different sensitivities of this function and for different values of θ. It can be seen that, for the more (less) sensitive $\tau(\cdot)$ function, the market under- (over-) exerts effort relative to the second-best. Remarks 1 and 2 suggest that the need for retention regulation is more prevalent in markets where the quality of originated assets is highly sensitive to originators' screening effort. If this is not the case, policy should focus on the problem of illiquidity in secondary markets by reducing retention.

Cross-subsidies play an important role in allowing the mechanism to implement different retention levels from those obtained in the market. To see this, I express the optimal mechanism $t = 1$ transfers as a function of the prices that would obtain in the market allocations for the same security:

$$t_b = E^b[X] + a^* \left[E^g_a[\min\{d^*, X\}] - E^b[X] + \frac{\pi}{\theta}(X_H - d^*)\right] \tag{33}$$

$$t_g = E^g_{a^*}[\min\{d^*, X\}] - \left(1 - a^*\right) \underbrace{\left[E^g_a[\min\{d^*, X\}] - E^b[X] + \frac{\pi}{\theta}(X_H - d^*)\right]}_{\text{Cross-Subsidy} \equiv \mathcal{T}_a(d^*)}, \tag{34}$$

where $\mathcal{T}_a(d^*_M) = 0$ since the incentive compatibility of the b-type binds at the market debt level. From inspecting the expression, we see that the transfer from the g-type to the b-type manager in $t = 1$ is positive when the mechanism implements less cash flow retention than the market, $d^* > d^*_M$, zero if cash flow retention is equal, $d^* = d^*_M$, and negative if the mechanism implements more retention, $d^* < d^*_M$. Therefore, cash flow retention rules should be accompanied by a subsidy to the senior claims issued by the g-type and a tax to the full claims issued by the b-type, while rules that reduce retention should be accompanied by a subsidy to the full claim of the b-type and a tax to senior claims of the g-type.

With two policy tools, a regulator can decentralize the optimal mechanism allocations. In particular, the regulator should use lump-sum transfers contingent on the issuance of different securities and a "skin-in-the-game" rule that imposes retention levels for certain issuances. The optimal regulation is summarized in the following corollary.

COROLLARY 3: *Second-best allocations are implemented in a setting with no commitment if the regulator offers the following two alternatives in secondary markets:*

1. *Sell a full claim to an asset cash flows and pay lump-sum tax* $T_J = a^*\mathcal{T}_{a^*}(d^*)$, *or*
2. *Sell a senior claim to an asset cash flows* $F(X) = \min\{d, X\}$, *retain the associated junior tranche, and receive lump-sum subsidy of* $T_S = (1 - a^*)\mathcal{T}_{a^*}(d^*)$,

where $\{a^*, d^*\}$ *are the second-best effort and debt levels.*

Policy plays a dual role in this environment. First, it reshapes incentives and improves upon the inefficiencies that arise due to lack of commitment. In some markets, regulation would improve liquidity at the expense of asset quality, while in others it would reduce market liquidity to improve the quality of originated assets. Second, and equally important, regulation provides stability to markets. I have shown that when the quality of assets is linked to the manager's screening choice (endogenous quality), the economy is prone to multiple equilibria when gains from trade are large, or an equilibrium may fail to exist when gains from trade are low. Regulation would therefore ensure that a unique equilibrium always exists.

A. Application: Regulating Markets for Securitized Assets

Regulators in the United States and Europe have implemented risk retention rules for all issuers of asset-backed securities (ABS). The rules require that all securitizers retain at least 5% risk exposure to the underlying cash flows of issued securities, with some exceptions. This intervention is commonly referred to as the "skin in the game" rule and is mandated in the Dodd–Frank Act for the United States and in the E.U. Capital Requirements Regulation (CRR) for Europe. These rules aim to address the misalignment of incentives between loan originators and investors, which are believed to have contributed to the financial crash of 2008. The model presented in this paper rationalizes the demand for cash flow retention rules as a way to provide incentives to originators to improve loan screening standards. However, the model suggests that demanding the same retention level of all issuers is inefficient. In particular, higher retention should be required of those issuers that claim to have good assets underlying their securities. The same retention level across issuances is not only costly in terms of gains from trade, but it makes it harder for those with good assets to signal their quality to the market, which in turn reduces incentives to screen loans. Furthermore, the model suggests that incentives are better provided when securitizers retain the first-loss piece of the underlying assets, while regulation allows issuers to choose the form of risk exposure (e.g., vertical slice, horizontal slice, originator's share, random selection of assets, or even exposure to assets that have the same underlying characteristics as the one backing the issued ABS).

In addition, the model suggests that retention rules should be accompanied by subsidies to the issuance of senior claims, which are financed by taxes to the issuance of equity (or junior) claims. To my knowledge, this type of policy has not been discussed in policy circles. Due to the presence of adverse selection in markets for securitized assets, transfers across issuers with different quality assets play an important role in implementing the desired retention levels. Thus, the model suggests that regulators should not only focus on retention levels for securitizers but also on the way the market compensates good versus bad issuers.

Finally, these retention regulations are accompanied by simplicity, disclosure, and transparency requirements and rules for originators on due diligence.[16] First, all information regarding the retention and risk exposure levels of originators/sponsors must be made available to investors. Second, investors and potential investors must have access to all material that is relevant to be able to assess the credit quality and performance of the assets underlying the issued securities and to all information that is necessary to perform stress tests

[16] See "Criteria for identifying simple, transparent and comparable securitizations," Basel Committee on Banking Supervision, http://www.bis.org/bcbs/publ/d332.pdf.

on the values of cash flows and collateral. It stands to reason that this type of regulation is beneficial when it is possible to implement. Giving investors access to all of the information required to evaluate underlying cash flows would solve both the moral hazard and the adverse selection problems. As I have shown, in a full-information environment, first-best allocations obtain.

V. Conclusions

This paper explores the tension between asset quality and liquidity in a model where an originator exerts effort to screen assets whose cash flows can be later sold in secondary markets. While screening improves asset quality, it introduces an asymmetric information problem that may reduce gains from trade with outside investors. The model is stylized, but has several predictions for the behavior of markets and highlights the need for policy intervention. I find that, under the optimal mechanism, costly retention of cash flows is essential to incentivize positive effort. In particular, imposing higher retention levels on those managers that hold better quality assets improves ex ante efficiency. In contrast, policies that require the same level of cash flow retention to all issuers of securities backed by previously originated assets can be too costly without necessarily improving incentives. Finally, the choice of screening effort under the optimal mechanism ensures that the social marginal benefit of asset screening is equal to its social marginal cost, which includes the costly retention required to implement a given level of effort.

The results in this paper contribute to the discussion on how to regulate markets for securitized assets. When markets lack commitment, equilibrium allocations can feature too much or too little effort relative to the second-best. This points to a need for policy intervention. In particular, some asset classes may feature excessive screening effort, resulting in inefficiently illiquid secondary markets, while others may feature too little screening effort, resulting in extremely liquid markets. All asset classes are not the same and hence regulation needs to be calibrated accordingly. Finally, when gains from trade are large, the economy becomes fragile and prone to multiple equilibria. Policy can implement second-best allocations with differential retention rules and transfers for different security issuances.

Initial submission: November 12, 2014; Accepted: September 19, 2016
Editors: Bruno Biais, Michael R. Roberts, and Kenneth J. Singleton

Appendix A: Endogenous Quality with Information Acquisition

In this section, I micro-fund the screening technology with endogenous quality presented in the main body of the paper. The main addition to the setup is that the manager has an information acquisition technology that is used to learn about potential projects and make financing decisions.

Let there be a continuum of positive net present value projects that pay high cash flows with probability $\pi > 0.5$. The manager can invest $C(a)$ in information to evaluate n projects chosen at random. It is essential that the screening capacity n be finite so that the result of the screening process is not always inferred by market investors in equilibrium. Investing $C(a)$ in information gives the manager access to a symmetric binary signal $s \in \{s_H, s_L\}$ for each project being evaluated, where:

$$\mathbb{P}_a\left(s_H|X_H\right) = \mathbb{P}_a\left(s_L|X_L\right) = \gamma(a), \tag{A.1}$$

where $a \mapsto \gamma(a)$, $\gamma(0) = \frac{1}{2}$, and $\gamma' > 0$. Thus, the unconditional probability of receiving the high signal is $\mathbb{P}_a(s_H) \equiv \pi\gamma(a) + (1-\pi)(1-\gamma(a))$ and is increasing in a for all $a \in (0, 1)$, and

$$\mathbb{P}_a\left(X_H|s_H\right) = \frac{\gamma(a)}{\mathbb{P}_a(s_H)}\pi \tag{A.2}$$

$$\mathbb{P}_a(X_L|s_L) = \frac{\gamma(a)}{1-\mathbb{P}_a(s_H)}(1-\pi). \tag{A.3}$$

Let $\tau(a) \equiv \mathbb{P}_a(X_H|s_H)$ for $a \in [0, 1]$. If the manager invests $a > 0$ and identifies a $s = s_H$ project, she will finance it since $\mathbb{P}_a(X_H|s_H) = \tau(a) > \pi > \mathbb{P}_a(X_H|s_L)$. If instead the manager does not observe any $s = s_H$ (she only observes $s = s_L$ for all evaluated projects), then she excludes the n projects that were evaluated and lends to a random project from the remaining pool, since $P(X_H|s_L) < \pi$. Finally, the probability of actually observing a high signal for at least one of the projects being evaluated, which I denote by $q(a)$, is given by

$$q(a) \equiv 1 - (1 - \mathbb{P}_a(s_H))^n. \tag{A.4}$$

Therefore, the probability of financing a high-quality project is given by

$$\mathbb{P}_a(X = X_H) = q(a)\tau(a) + (1 - q(a))\pi, \tag{A.5}$$

which is increasing in a. In addition, we have that $\tau(0) = \pi$, $\tau'(\cdot) > 0$, $q'(\cdot) > 0$, $\forall a \in [0, 1]$, and $\mathbb{P}_0(X = X_H) = \pi$. Thus, for a given screening capacity n, this framework generates the screening structure in Figure A1.

As in the Endogenous Quality Case, investment in information (i.e., effort) increases both the likelihood of becoming a good-type manager and the expected

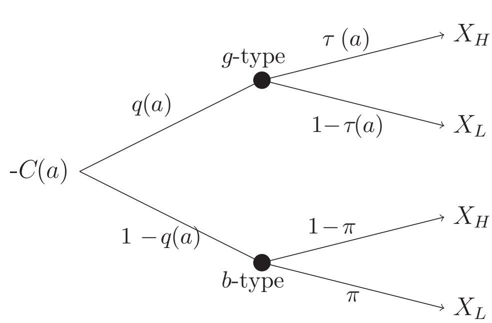

Figure A1. Project screening with an information acquisition technology.

cash flow conditional on being a good-type manager. Furthermore, it decreases the likelihood of investing in the "not-so-good" project, whose expected cash flow is equivalent to that obtained if there is no project screening.

Appendix B: Endogenous versus Exogenous Quality

Proposition 2 in Section II characterizes the solution to the optimal mechanism as a function of the marginal benefits and costs of effort. The following table depicts the difference between the marginal return to effort and indirect cost in the exogenous versus endogenous quality cases,

	Endo : $\tau(0) = \pi$ and $\tau'(\cdot) > 0$	**Exo** : $\tau(a) = \bar{\tau} > \pi, \forall a$
$\rho'(a)$	$\tau(a) - \pi + a\tau'(a)$	$\bar{\tau} - \pi$
$C_R(a)$	$(\theta - 1)\frac{\tau(a)}{\rho'(a)} aC'(a)$	$(\theta - 1)\frac{\bar{\tau}}{\rho'(a)} aC'(a)$
$C_R'(a)$	$(\theta - 1)\frac{\tau(a)}{\rho'(a)}[\mathcal{J}(a)C'(a) + aC''(a)]$	$(\theta - 1)\frac{\bar{\tau}}{\rho'(a)}[C'(a) + aC''(a)]$

where $\mathcal{J}(a) \equiv 1 + \frac{a\tau'(a)}{\tau(a)} - \frac{a\rho''(a)}{\rho'(a)}$. For a given quality level (i.e., for a such that $\tau(a) = \bar{\tau}$), the marginal return to effort is higher when quality is endogenous. As a result, the indirect cost of effort is lower, since effort is easier to implement when its marginal return is higher. The relation between marginal indirect costs, however, depends on the relation between $\tau(\cdot)$ and $\rho'(\cdot)$, as captured by $\mathcal{J}(\cdot)$.

Appendix C: Proofs of Section I

PROOF OF PROPOSITION 1: In $t = 1$, the manager chooses what cash flows to sell. With full information, the value for the manager in $t = 1$ is

$$V_1(z) = \theta p(F) + \mathbb{E}_a^z[X - F(X)] = (\theta - 1)\mathbb{E}_a^z[F(X)] + \mathbb{E}_a^z[X], \tag{C.1}$$

where the last equality arises since competitive investors will price any security at its expected value, which is the same for investors as for the manager. Given the feasibility constraint, the manager chooses to sell a full claim to her cash flows, $F_{FB}^*(X) = X$, independent of her z-type and her initial choice of effort a. Since the manager values funds in $t = 1$, selling X implements allocative efficiency.

Screening effort is chosen to maximize the value for the manager in $t = 0$:

$$V_0(a, a) = \theta E_a[X] - C(a), \tag{C.2}$$

where $E_a[X] = \rho(a)X_H + (1 - \rho(a))X_L$. By Assumption 1, the first-order condition given in (5) characterizes the solution to the problem $\max_a V_0(a, a)$ both in the first-best and in the full information equilibrium.

Appendix D: Proofs of Section II

A. Proofs of Lemmas and Propositions in Section II

All complementary lemmas used in this section are in Appendix B.

PROOF OF LEMMA 1: Lemmas 4 and 5 show that under the optimal mechanism the participation constraint (PC) of investors and the incentive compatibility constraint (IC) of the b-type bind. By plugging in the binding PC of investors to the value for the manager in $t = 0$, we obtain

$$V_0(a) = \theta E_a[X] - (\theta - 1)[aE_a^g[T_g(X)] + (1 - a)E^b[T_b(X)]] - C(a). \quad \text{(D.1)}$$

And by plugging in the binding IC for the b-type into the IC for effort, we obtain (see Lemma 6)

$$\rho'(a^*)(T_g(X_H) - T_g(X_L)) = C'(a^*). \quad \text{(D.2)}$$

Therefore, an optimal mechanism $\{a^*, t_z, T_z(\cdot)\}$ maximizes (D.1) subject to (D.2) and to the IC for type revelation and the PC of investors. I begin by showing that $t = 2$ transfers for the b-type are zero. The idea for the proof is as follows. If the b-type is retaining cash flows, it is optimal to sell some of these cash flows and split the surplus equally among all manager types. This adjustment does not tighten the constraints and is profitable due to gains from trade.

(a) *b-type t=2 transfers*. Assume that $T_b(X_H) > 0$ under the optimal mechanism. Next, reduce the $t = 2$ transfer for the b-type, $T_b'(X_H) = T_b(X_H) - \epsilon$, by $\epsilon > 0$ small so that $T_b'(X_H) > 0$, and increase her $t = 1$ transfer so that she is indifferent, $t_b' = t_b + \frac{\pi}{\theta}\epsilon$. Then the IC of the b-type continues to bind with the new transfers. Note that the IC of the g-type is relaxed since $\tau(a) \geq \pi$ for any a, and the IC constraint for effort (D.2) is unaffected. However, the PC of investors is relaxed. Let us now transfer the extra surplus, $(1 - a)(\pi\epsilon - \frac{\pi}{\theta}\epsilon) = (1 - a)\frac{\theta - 1}{\theta}\pi\epsilon > 0$, equally to both manager types. This increases the value of the manager (D.1) and does not change any of the constraints. Contradiction. Thus, $T_b(X_H) = 0$, which by feasibility implies that $T(X_L) = 0$, that is, $T_b(X) = 0$.
Note that the binding IC for the b-type implies an always-slack IC for the g-type as $T_b(\cdot) = 0$: $\theta t_g + E^b[T_g(X)] = \theta t_b < \theta t_g + E^g[T_g(X)]$, since $E^b[T_g(X)] < E^g[T_g(X)]$ for any $T_g \in \Delta$. Therefore, $t = 1$ transfers are chosen to guarantee that both the IC for type revelation and PC of investors hold for any choice of $t = 2$ transfers and implementable effort levels. Given this, I am left to characterize the $t = 2$ transfers for the g-type that maximize (D.1) subject to (D.2). I do so in two parts. Part 1. I show that $T_g(X) = \max\{0, X - d\}$ for some $d \in [0, X_H]$: it is optimal to transfer junior claims to the g-type since this incentivizes effort at the lowest

retention cost. Part 2. I show that $d \in [X_L, X_H]$: there are no gains from imposing retention in the low cash flow state.

(b) *g-type t=2 transfers*. Part 1. Assume that $T_g(X) \neq \max\{0, X-d\}$ for any $d \in [0, X_H]$. This means that there exists $f_H \in [0, X_H]$ and $f_L \in (0, X_L]$ such that $T_g(X_H) = X_H - f_H$ and $T_g(X_L) = X_L - f_L$, where $f_H - f_L > 0$ due to monotonicity. Define new transfer function $G(X) = \max\{0, X-d\}$ with $d \in [0, X_H]$ chosen so that the expected value of the $t = 2$ transfers to the g-type remains unchanged:

$$E^g_{a^*}[T_g(X)] = E^g_{a^*}[G(X)] \tag{D.3}$$

$$\tau(a^*)(T_g(X_H) - T_g(X_L)) + T_g(X_L) = \tau(a^*)(G(X_H) - G(X_L)) + G(X_L)\tau(a^*)$$

$$(X_H - X_L - (f_H - f_L)) + X_L - f_L = \begin{cases} \tau(a^*)(X_H - d) & d \geq X_L \\ \tau(a^*)(X_H - X_L) + X_L - d & d < X_L, \end{cases} \tag{D.4}$$

where I have used the fact that $G(X_L) = 0$ if $d \geq X_L$ and $G(X_L) = X_L - d$ otherwise. Note that, for any d, $T_g(X_H) - T_g(X_L) < G(X_H) - G(X_L)$. Therefore, the $t = 2$ transfers given by G implement higher effort:

$$\rho'(a^*)(G(X_H) - G(X_L)) - C'(a^*) > \rho'(a^*)(T_g(X_H) - T_g(X_L)) - C'(a^*) = 0. \tag{D.5}$$

Now, to implement the same level of effort a^* as before, define new transfers $G'(X) = \max\{0, X - d'\}$ so that the following first-order condition holds:

$$\rho'(a^*)(X_H - d' - \max\{0, X_L - d'\}) - C'(a^*) = 0. \tag{D.6}$$

Note that $d' > d$, and therefore $E^g_{a^*}[G'(X)] < E^g_{a^*}[T_g(X)]$. Hence, transfers given by G' implement the same level of effort a^* and generate higher manager value in $t = 0$ by (D.1) since they imply less cash flow retention. Contradiction. Since $T_g(X)$ was an arbitrary transfer function in Δ, in the optimal mechanism $T_g(X) = \max\{0, X-d\}$ for some $d \in [0, X_H]$.

Part 2. In this part, I show that the retention of cash flows in the bad cash-flow realization do not improve incentives and are costly since they reduce gains from trade. Suppose that $T_g(X) = \max\{0, X-d\}$, with $d < X_L$. Define a new $t = 2$ transfer function by $H(X) = \max\{0, X - X_L\}$; that is, $H(X_H) = X_H - X_L$ and $H(X_L) = 0$. Then $H(X_H) - H(X_L) = X_H - X_L = T_g(X_H) - T_g(X_L)$, so the transfers given by H implement the same level of effort a^* as transfers given by T_g. However, note that

$$E^g_{a^*}[H(X)] = \tau(a^*)(X_H - X_L) < \tau(a^*)(X_H - X_L) + X_L - d = E^g_{a^*}[T_g(X)], \tag{D.7}$$

and thus H generates higher manager value in $t = 0$ (see D.1). Contradiction. Therefore, it must be the case that under the optimal mechanism, $T_g(X_L) = 0$ and $T_g(X_H) = X_H - d$ for $d \in [X_L, X_H]$.

PROOF OF LEMMA 2: Lemmas 4 and 5 show that under the optimal mechanism the PC of investors and the b-type IC bind. The $t = 1$ transfers are pinned down by the binding constraints:

$$a t_g + (1 - a) t_b = a E_a^g[\min\{d, X\}] + (1 - a) E^b[X] \tag{D.8}$$

$$\theta(t_b - t_g) = E^b[\max\{0, X - d\}], \tag{D.9}$$

where results from Lemma 1 have been incorporated. By solving this system of equations for $\{t_g, t_b\}$, the expressions in (12) are obtained.

PROOF OF COROLLARY 1: From Lemmas 4 and 5 we know that the PC of investors and the IC for the b-type manager bind. From Lemma 1, we know that $T_b(\cdot) = 0$, $T_g(X_L) = 0$, and $T_g(X_H) = X_H - d$ for $d \in [X_L, X_H]$.

Plugging the binding PC of investors and the value of the $t = 2$ transfers, the value to the manager in $t = 0$ for a given effort choice a and debt level d is

$$\theta E_a[X] - (\theta - 1) a \tau(a)(X_H - d) - C(a). \tag{D.10}$$

From Lemma 6, the IC constraint for effort is reduced to

$$\rho'(a)(X_H - d) = C'(a) \tag{D.11}$$

by plugging in the binding IC of the b-type and the results from Lemmas 1 and 2. Therefore, the mechanism chooses $\{a, d\} \in [0, 1] \times [X_L, X_H]$ to maximize (D.10) subject to (D.11), since $t = 1$ transfers, $\{t_g^*, t_b^*\}$, ensure that the IC for type revelation and the PC of investors hold for any given pair $\{a^*, d^*\}$, and where a binding IC for the b-type ensures a slack IC for the g-type.

PROOF OF PROPOSITION 2: The allocations are obtained by solving the problem from Corollary 1, where by Assumption 1 the corner $a = 0$ has been ruled out and first-order conditions characterize the interior solution. When the constraints $X_L \leq d \leq X_H$ do not bind, we have an interior solution given by the first-order condition with respect to a and the constraint:

$$a^* : \quad \theta\rho'(a^*)(X_H - X_L) - C'(a^*) - C_R'(a^*) = 0 \tag{D.12}$$

$$d^* : \quad \rho'(a^*)(X_H - d^*) - C'(a^*) = 0. \tag{D.13}$$

Note that $d^* \leq X_H$ and thus the upper bound constraint never binds. However, if $\frac{C'(a^*)}{\rho'(a^*)} > X_H - X_L$ for the interior level a^*, then we have a corner solution where $d^* = X_L$ and a^* is given by $X_H - X_L = \frac{C'(a^*)}{\rho'(a^*)}$. Finally, Lemma 7 shows that the solution to the above problem is robust to global deviations.

PROOF OF PROPOSITION 3: Define functions $H(\cdot; \omega)$ and $G(\cdot, \cdot; \omega)$, where $\omega \equiv (\theta, \chi, X_H - X_L)$, by

$$H(a; \omega) \equiv \theta\rho'(a)(X_H - X_L) - C'(a) - C_R'(a) \tag{D.14}$$

$$G(a,d;\omega) \equiv \rho'(a)(X_H - d) - C'(a), \tag{D.15}$$

and note that, at the optimal mechanism allocations $\{a^*, d^*\}$, $H(a^*,\omega) = 0$ and $G(a^*, d^*;\omega) = 0$ by Proposition 2.

Gains from Trade: By the second-order conditions of the optimal mechanism problem (Assumption 1, (iii)), $H_a(a^*;\omega) < 0$. Since

$$C_R'(a) = (\theta - 1)\tau(a)\frac{C'(a)}{\rho'(a)}\left[1 + \frac{a\tau'(a)}{\tau(a)} - \frac{a\rho''(a)}{\rho'(a)} + \frac{aC''(a)}{C'(a)}\right], \tag{D.16}$$

we have that

$$H_\theta(a^*,\omega) = \rho'(a^*)(X_H - X_L) - \frac{C_R'(a^*)}{\theta - 1} \tag{D.17}$$

$$= \frac{1}{\theta - 1}[C'(a^*) - \rho'(a^*)(X_H - X_L)] < 0, \tag{D.18}$$

where the inequality is strict since $d^* > X_L$ and $C'(a^*) = \rho'(a^*)(X_H - d^*)$. Thus, $\frac{da^*}{d\theta} = -\frac{H_\theta(a^*;\omega)}{H_a(a^*;\omega)} < 0$.

By the second-order conditions of the manager's problem (Assumption 1, (ii)), we know that $G_a(a^*, d^*;\omega) < 0$ and $G_d(a^*, d^*;\omega) < 0$, and in addition $G_\theta(a^*, d^*, \omega) = 0$. Thus, it follows that $\frac{dd^*}{d\theta} > 0$ since $\frac{da^*}{d\theta} < 0$.

Cost of Effort. Since $C(a) = \chi h(a)$, we have

$$C_R'(a) = (\theta - 1)\tau(a)\frac{\chi h'(a)}{\rho'(a)}\left[1 + \frac{a\tau'(a)}{\tau(a)} - \frac{a\rho''(a)}{\rho'(a)} + \frac{ah''(a)}{h'(a)}\right]. \tag{D.19}$$

We also have that $H_\chi(a^*;\omega) = -h'(a^*) - \frac{C_R'(a^*)}{\chi} < 0$. Thus, $\frac{da^*}{d\chi} = -\frac{H_\chi(a^*;\omega)}{H_a(a^*;\omega)} < 0$.

On the other hand, the effect on debt level d^* is given by

$$\frac{dd^*}{d\chi} = -\frac{h'(a^*)}{\rho'(a^*)} - \chi\frac{d}{da^*}\left(\frac{h'(a^*)}{\rho'(a^*)}\right)\left(\frac{da^*}{d\chi}\right) \tag{D.20}$$

$$= -\frac{h'(a^*)}{\rho'(a^*)}\left(1 + \frac{da^*}{d\chi}\frac{\chi}{a^*}\left(\frac{a^*h''(a^*)}{h'(a^*)} - \frac{a^*\rho''(a^*)}{\rho'(a^*)}\right)\right). \tag{D.21}$$

Therefore, $\frac{dd^*}{d\chi} < 0$ if and only if $\chi[\frac{\rho''(a^*)}{\rho'(a^*)} - \frac{h''(a^*)}{h'(a^*)}]\frac{da^*}{d\chi} < 1$, where $[\frac{\rho''(a^*)}{\rho'(a^*)} - \frac{h''(a^*)}{h'(a^*)}] < 0$ by Assumption 1.

Differential Quality. From the equation $H(a^*;\omega) = 0$ and the second order conditions, it follows immediately that $\frac{da^*}{d(X_H - X_L)} > 0$. Analogously, from the equation $G(a^*, d^*;\omega) = 0$, since $\frac{da^*}{d(X_H - X_L)} > 0$, it follows that $\frac{d(X_H - d^*)}{d(X_H - X_L)} > 0$.

B. Complementary Lemmas for Section II

LEMMA 3: *In the exogenous quality case, the IC for effort (10) can be replaced by*

$$a = \arg\max_{\hat{a}} \hat{a}(\theta t_g + E^g[T_g(X)]) + (1-\hat{a})(\theta t_b + E^b[T_b(X)]) - C(\hat{a}). \quad \text{(D.22)}$$

In the endogenous quality case, the IC for effort (10) can be replaced by the following two constraints:

$$a = \arg\max_{\hat{a}} \hat{a}\left(\theta t_g + E^g_{\hat{a}}[T_g(X)]\right) + (1-\hat{a})\left(\theta t_b + E^b[T_b(X)]\right) - C(\hat{a}) \quad \text{(D.23)}$$

$$\max_{\hat{a}} \hat{a}\left(\theta t_g + E^g_{\hat{a}}[T_g(X)]\right) + (1-\hat{a})\left(\theta t_b + E^b[T_b(X)]\right) - C(\hat{a}) \geq$$
$$\max_{\hat{a}} \theta t_b + \hat{a}E^g_{\hat{a}}[T_b(X)] + (1-\hat{a})E^b[T_b(X)] - C(\hat{a}). \quad \text{(D.24)}$$

PROOF: First, note that the IC for the b-type is independent of a, and thus holds for all $a \in [0, 1]$. In the exogenous-quality case, this is also true for the g-type IC. Thus, in the exogenous-quality case, we can rewrite the IC for effort using the fact that both ICs for type revelation hold, which leads to the expression in the lemma.

When quality is endogenous, however, the IC for the g-type does depend on a. Therefore, to ensure that the mechanism is robust to global deviations, the constraint (D.24) is imposed. The constraint ensures that reporting a g-type truthfully and exerting the corresponding best-response effort gives the manager at least as much value as deviating both on her effort choice and on her ex post report of type.

LEMMA 4: *Under the optimal mechanism, the PC of investors binds.*

PROOF: Assume not. If the PC of investors is slack, we can define new $t = 1$ transfers $t'_z = t_z + \epsilon$ for $z \in \{g, b\}$ where $\epsilon > 0$ is small enough that the investors' PC continues to hold. These new transfers increase the value to the manager and do not affect the other constraints. Contradiction.

LEMMA 5: *Under the optimal mechanism, the IC constraint of the b-type binds.*

PROOF: Assume that, under the optimal mechanism $\{a^*, t_g, t_b, T_g(\cdot), T_b(\cdot)\}$, the IC for the b-type does not bind.

First, note that if $T_g(\cdot) = T_b(\cdot) = 0$ or if $a^* = 0$, then both ICs for type revelation bind. Thus, it must be that $a^* > 0$ and that either $T_g(\cdot) \neq 0$ or $T_b(\cdot) \neq 0$. Suppose $T_g(\cdot) = 0$. Combining the two ICs we get $E^g_{a^*}[T_b(X)] \leq \theta(t_g - t_b) < E^b[T_b(X)]$. But this is not possible due to monotonicity and the fact that $\tau(a^*) > \pi$. Thus, $T_g(\cdot) \neq 0$, which implies that $T_g(X_H) > 0$.

Given the slack IC for the b-type, transfer a small amount from the b-type to the g-type in $t = 1$: $t'_b = t_b - a\epsilon$ and $t'_g = t_g + \epsilon(1-a)$ for $\epsilon > 0$ sufficiently small that the IC of the b-type remains slack. Note that the PC of investors is not affected but the differential between $t = 1$ transfers has increased by ϵ: $t'_g - t'_b = t_g - t_b + \epsilon$. Now decrease the $t = 2$ transfer to the g-type in the high cash flow state (since I have shown it is positive) by ε: $T'_g(X_H) = T_g(X_H) - \varepsilon$ for $\varepsilon > 0$ sufficiently small that $T'_g(X_H) > 0$. Choose pair $\{\epsilon, \varepsilon\}$ so that the IC for effort (11) holds at the optimal mechanism effort level a^* when evaluated with the new transfers:

$$\begin{aligned}\theta\left(t_g - t_b + \epsilon\right) + \left(\tau(a^*) + a^*\tau'(a^*)\right)\left(T_g(X_H) - T_g(X_L) - \varepsilon\right)\\ -\pi\left(T_b(X_H) - T_b(X_L)\right) + T_g(X_L) - T_b(X_L) - C'(a^*) = 0.\end{aligned}$$

Such a pair exists and satisfies

$$\theta\epsilon = \varepsilon\left(\tau(a^*) + a^*\tau'(a^*)\right).$$

Note that the IC for the g-type is relaxed since the increase in $t = 1$ transfers more than compensates for the decrease in the $t = 2$ transfer in the high state: $\theta\epsilon - \tau(a^*)\varepsilon \geq 0$. In addition, this decrease in $t = 2$ transfers generates an extra surplus: $a^*\tau(a^*)\varepsilon$ (i.e., the PC for investors is now slack, since reducing this transfer is equivalent to allowing the manager to sell more to investors). This surplus can now be used to increase $t = 1$ transfers of both manager types by $a^*\tau(a^*)\varepsilon$, that is, until the PC of investors binds again. Thus, we can rewrite the objective function of the manager in $t = 0$ as follows:

$$\theta E_a[X] - (\theta - 1)\left[aE^b_a\left[T_b(X)\right] - a\tau(a)\varepsilon + (1 - a)E^b[T_b(X)]\right] - C(a), \quad \text{(D.25)}$$

where $E_a[X] \equiv \rho(a)X_H + (1 - \rho(a))X_L$. Therefore, the new transfers implement the same effort level a^*, and increase the value of the manager for $\varepsilon > 0$, without affecting any of the other constraints. Contradiction.

LEMMA 6: *The IC for effort under the optimal mechanism can be rewritten as*

$$\rho'(a)(T_g(X_H) - T_g(X_L)) = C'(a). \quad \text{(D.26)}$$

PROOF: The binding IC of the b-type (by Lemma 5) imposes a condition on the difference between the $t = 1$ transfers:

$$\begin{aligned}\theta\left(t_g - t_b\right) &= E^b\left[T_b(X) - T_g(X)\right]\\ &= \pi(T_b(X_H) - T_g(X_H)) + (1 - \pi)(T_b(X_L) - T_g(X_L)).\end{aligned}$$

Plugging the binding IC of the b-type into the IC for effort resulting from the first-order approach (11), we obtain

$$\left(\tau(a) - \pi + a\tau'(a)\right)\left(T_g(X_H) - T_g(X_L)\right) - C'(a) = 0, \quad \text{(D.27)}$$

where the expression is obtained since $\rho'(a) = \tau(a) - \pi + a\tau'(a)$.

LEMMA 7: *There are no profitable global deviations from the solution characterized by the first-order approach.*

PROOF: Proposition 2 presents the solution to the optimal mechanism obtained with the use of the first-order approach for the IC for effort. From Lemma 3, we know that, for the exogenous quality case, the solution always satisfies global incentive compatibility. However, also by Lemma 3, when quality is endogenous, there is an additional constraint that we need to verify to rule out global deviations. In what follows, I verify that the constraint for global incentive compatibility holds. Plugging the results from Lemmas 1 and 2, into the no-global deviations constraint for the endogenous-quality case from Lemma 3, we obtain

$$\max_{\hat{a}\in[0,1]} \hat{a}(\theta t_g + E^g_{\hat{a}}[\max\{0, X-d\}]) + (1-\hat{a})\theta t_b - C(\hat{a}) \geq \max_{\hat{a}\in[0,1]} \theta t_b - C(\hat{a}) \tag{D.28}$$

$$\max_{\hat{a}\in[0,1]} \hat{a}\left(\theta(t_g - t_b) + E^g_{\hat{a}}[\max\{0, X-d\}]\right) - C(\hat{a}) + \theta t_b \geq \theta t_b, \tag{D.29}$$

which always holds since

$$\max_{\hat{a}\in[0,1]} \hat{a}\left(\theta(t_g - t_b) + E^g_{\hat{a}}[\max\{0, X-d\}]\right) - C'(\hat{a}) \geq 0 \tag{D.30}$$

for any $d \in [X_L, X_H]$, as $a^* = 0$ is a possible solution to the maximization problem on the left-hand side.

Appendix E: Proofs of Section III

A. Proofs of Lemmas and Propositions in Section III

PROOF OF PROPOSITION 4: Let $\{F^*_b, F^*_g\} \in \Delta^2$ be the securities sold in equilibrium in secondary markets by the b- and g-type manager, respectively. By Lemma 8, the equilibrium is separating (pooling equilibria do not survive D1-Refinements). Therefore, the b-type obtains its full-information payoff $u^*_b \equiv \theta E^b[X]$, since $\mu(F_b) = 0$ and $F^*_b(X) = X$. By Lemma 9, F^*_g has to solve the following problem:

$$\max_{F\in\Delta} \quad \theta E^g_{a^e}[F(X)] + E^g_a[X - F(X)] \tag{E.1}$$

$$\text{s.t.} \quad \theta E^g_{a^e}[F(X)] + E^b[X - F(X)] = u^*_b. \tag{E.2}$$

That is, security $F_g \in \Delta$ maximizes the value of the g-type manager in $t = 1$, subject to the b-type being indifferent between mimicking or receiving her full-information payoff. The problem can be rewritten by subtracting u^*_b from the objective, to obtain

$$\max_{F\in\Delta} \quad (\tau(a) - \pi)[X_H - X_L - (F(X_H) - F(X_L))], \tag{E.3}$$

subject to the b-type having value u_b^*

$$(\theta\tau(a^e)-\pi)(F(X_H)-F(X_L))+(\theta-1)F(X_L)+E^b[X]=u_b^*, \tag{E.4}$$

where for any $\{a, a^e\}$, $\pi \leq \tau(a)$ and $\pi \leq \theta\tau(a^e)$.

Part 1. I show that $F_g(X_L)=X_L$. Assume that $F_g(X_L)<X_L$. Define a new security with $F_g'(X_L)=F_g(X_L)+\epsilon$ and $F_g'(X_H)=F_g'(X_H)-\varepsilon$, where $\epsilon>0$, $\varepsilon>0$ and both are sufficiently small that $F_g'\in\Delta$ and the b-type value is unchanged (i.e., constraint (E.4) holds):

$$(\theta\tau(a^e)-\pi)(-\varepsilon-\epsilon)+(\theta-1)\epsilon=0. \tag{E.5}$$

Finally, note that $F_g'(X_H)-F_g'(X_L)=F_g(X_H)-F_g(X_L)-\frac{(\theta-1)\epsilon}{\tau(a^e)-\pi}$ and thus the new security increases the objective (E.3). We therefore have $F_g(X_L)=X_L$, which combined with monotonicity implies that we can express the security sold by the g-type as a debt-like security with debt level greater than or equal to X_L: $F_g(X)=\min\{d, X\}$ for $d\in[X_L, X_H]$.

Part 2. I show that, when $\theta\tau(a^e)\geq\tau(a)$, the debt level is chosen to make the IC of the b-type bind. Otherwise, the g-type manager sells a risk-free claim.

Let debt level d^{IC} be given by the binding IC of the b-type manager:

$$\theta E^g_{a^e}[F_g(X)]+E^b[X-F_g(X)]=\theta E^b[X] \tag{E.6}$$

$$\Rightarrow (\theta\tau(a^e)-\pi)(d^{IC}-X_L)=(\theta-1)\pi(X_H-X_L). \tag{E.7}$$

Consider $d>d^{IC}$. Note that d violates the IC of the b-type. Therefore, there is pooling in secondary markets, which by Lemma 8 cannot be an equilibrium.

Consider $d<d^{IC}$.

(a) When $\theta\tau(a^e)-\tau(a)>0$, there is a profitable deviation to issue debt with debt level: $d'\in(d, d^{IC})$. To see this, note that from the IC of the b-type, the set of beliefs for which the b-type benefits from deviating to debt with d' is empty, since the IC is slack for debt levels below d^{IC} for $\mu\in[0,1]$. The g-type's extra payoff from deviating is given by $(\theta\tau(a^e)-\tau(a))(d'-d)$ and this deviation is strictly profitable for $\mu\in(\underline{\mu}, 1]$ when $\theta\tau(a^e)-\tau(a)>0$. Therefore, the belief assigned to this deviation is $\mu=1$. As a result, the LCSE with $F_g(X_L)=\min\{d^{IC}, X\}$ and $F_b(X)=X$ is the unique equilibrium in secondary markets.

(b) When $\theta\tau(a^e)=\tau(a)$, it is without loss of generality to assume that $d=d^{IC}$ since the g-type is indifferent between selling any debt level above X_L.

(c) When $\theta\tau(a^e)<\tau(a)$, the discount received in the market by the g-type is large enough that she prefers to sell a risk-free claim, $F_g(X)=X_L$, while the IC of the b-type is slack when $d=X_L$: the b-type is strictly better off selling equity. In addition, to price this risk-free claim investors do not form beliefs, and thus no beliefs are assigned to this deviation from the LCSE strategies. Therefore, the unique equilibrium in secondary markets has $F_b(X)=X$ and $F_g(X)=X_L$.

PROOF OF PROPOSITION 5: In any equilibrium, $a = a^e$. Using the results from Proposition 4, the problem of the manager in $t = 0$ for market beliefs $a^e \in [0, 1]$ can be written as

$$\max_{a \in [0,1]} a[\max\{\theta\tau(a^e) - \tau(a), 0\}(d(a^e) - X_L)] + (1-a)(\theta - 1)\pi(X_H - X_L)$$
$$+E_a[X] - C(a), \tag{E.8}$$

where $d(a^e) = X_L + \frac{(\theta-1)\pi}{(\theta\tau(a^e)-\pi)}(X_H - X_L)$. Note that the objective is differentiable with respect to a at $a = a^e$, since $\theta\tau(a^e) > \tau(a) = \tau(a^e)$. Thus, in any equilibrium, a^* has to satisfy the first-order condition

$$\left(1 - \frac{(\theta-1)\pi}{\theta\tau(a^*) - \pi}\right)\rho'(a^*)(X_H - X_L) - C'(a^*) = 0, \tag{E.9}$$

where I first differentiate the objective function with respect to a around a^e and I then impose the equilibrium condition $a^* = a^e$. Finally, it follows from Proposition 4 that, when $a^* = a^e$, debt level d^* has to satisfy:

$$d^* = X_L + \frac{(\theta-1)\pi}{\theta\tau(a^*) - \pi}(X_H - X_L). \tag{E.10}$$

[Endogenous Quality:] There are at least two solutions to the system of equations (E.9)–(E.10). First, the corner solution with zero effort and no retention is a solution. This is because no effort is consistent with no retention, and vice-versa: $C'(0) = 0$ and $(1 - \frac{(\theta-1)\pi}{\theta\tau(0)-\pi}) = 0$ since $\tau(0) = \pi$. Second, there is an interior solution with positive effort level $a^*_M > 0$, guaranteed by Assumption 1, and positive retention $d^*_M < X_H$, given by (E.10). These two are the only two equilibrium candidates for the endogenous-quality case whenever $\frac{C'(a)}{\rho'(a)}\frac{\theta\tau(a)-\pi}{\theta(\tau(a)-\pi)}$ is increasing in a.

[Exogenous Quality:] The first-order condition that determines effort becomes:

$$\left(1 - \frac{(\theta-1)\pi}{\theta\bar{\tau} - \pi}\right)(\bar{\tau} - \pi)(X_H - X_L) - C'\left(a^*_M\right) = 0. \tag{E.11}$$

Thus, there is a unique a^*_M that solves this problem, and a unique d^*_M given by (E.10). This is because the retention in secondary markets is independent of the manager's effort choice and of market beliefs about this action. Thus, there is only one candidate equilibrium.

PROOF OF PROPOSITION 6: Define functions $H(\cdot;\omega)$ and $G(\cdot,\cdot;\omega)$, where $\omega = (\theta, \chi, X_H - X_L)$ by

$$H(a;\omega) \equiv \rho'(a)\frac{\theta(\tau(a) - \pi)}{\theta\tau(a) - \pi}(X_H - X_L) - C'(a) \tag{E.12}$$

$$G(a, d;\omega) \equiv X_L + \frac{(\theta-1)\pi}{\theta\tau(a) - \pi}(X_H - X_L) - d, \tag{E.13}$$

and note that, at the equilibrium allocation with positive effort $\{a_M^*, d_M^*\}$, $H(a_M^*;\omega) = 0$ and $G(a_M^*, d_M^*;\omega) = 0$ by Proposition 5. Recall also that I assume that $\frac{C'(a)}{\rho'(a)}\frac{\theta\tau(a)-\pi}{\theta(\tau(a)-\pi)}$ is increasing in a, which ensured that there is at most one positive effort equilibrium (see proof of Proposition 5). This in turn implies that $H_a(a_M^*;\omega) < 0$.

Gains from Trade. Since we have that $H_\theta(a_M^*;\omega) < 0$ and $H_a(a_M^*;\omega) < 0$, it follows that $\frac{da_M^*}{d\theta} < 0$.

For the debt level, since $G_a(a_M^*, d_M^*;\omega) < 0$, $G_\theta(a_M^*, d_M^*;\omega) > 0$, and $G_d(a_M^*, d_M^*;\omega) < 0$, it follows that

$$\frac{dd_M^*}{d\theta} = -\frac{G_a(a_M^*, d_M^*;\omega)\frac{da_M^*}{d\theta} + G_\theta(a_M^*, d_M^*;\omega)}{G_d(a_M^*, d_M^*;\omega)} > 0. \tag{E.14}$$

Cost of Effort. Note that $H_\chi(a_M^*;\omega) < 0$, and because $H_a(a_M^*;\omega) < 0$,

$$\frac{da_M^*}{d\chi} = -\frac{H_\chi(a_M^*;\omega)}{H_a(a_M^*;\omega)} < 0. \tag{E.15}$$

It follows immediately from $G(a_M^*, d_M^*;\omega) = 0$ that $\frac{dd_M^*}{d\chi} \geq 0$. In particular, the inequality is strict in the endogenous-quality case, where $\tau'(\cdot) > 0$. In contrast, when quality is exogenous, $\tau(a) = \bar{\tau}$ is constant. Since the cost function does not directly affect debt levels, the latter are unaffected.

Differential Quality. Note that $H_{(X_H - X_L)}(a_M^*;\omega) > 0$, and because $H_a(a_M^*;\omega) < 0$

$$\frac{da_M^*}{d(X_H - X_L)} = -\frac{H_{(X_H - X_L)}(a_M^*;\omega)}{H_a\left(a_M^*;\omega\right)} > 0. \tag{E.16}$$

The equation $G(a_M^*, d_M^*;\omega) = 0$ implies that

$$X_H - d_M^* = \frac{\theta\left(\tau(a_M^*) - \pi\right)}{\theta\tau\left(a_M^*\right) - \pi}(X_H - X_L), \tag{E.17}$$

which is weakly increasing in $X_H - X_L$, since $\tau'(\cdot) \geq 0$ and $\frac{da_M^*}{d(X_H - X_L)} > 0$. Thus, $X_H - d_M^*$ is increasing in $X_H - X_L$.

PROOF OF PROPOSITION 7: [Endogenous Quality:] It remains to check when, and if, the candidates from Proposition 5: $\{0, X_H\}$ and $\{a_M^*, d_M^*\}$ given by (21) and (22) are equilibria. First, note that deviations to lower effort levels do not change outcomes in secondary markets (LCSE outcomes) and thus are not optimal ex ante. Therefore, only deviations to higher effort levels, $a > a^*$, need to be ruled out. For this type of deviation to be possible, it must be that the manager deviates to $a > a^*$ such that $\theta\tau(a^*) = \theta\tau(a^e) < \tau(a)$, and later deviates to issue a risk-free claim. The best deviation on effort is therefore given by $\{\tilde{a}, X_L\}$, where

$$\tilde{a} \equiv \arg\max_{a\in[0,1]}(\theta - 1)\left((1 - a)\pi(X_H - X_L) + X_L\right) + E_a[X] - C(a). \tag{E.18}$$

I proceed to characterize the cases in which deviations to $\{\tilde{a}, X_L\}$ from our equilibrium candidates $\{0, X_H\}$ and $\{a_M^*, d_M^*\}$ are not profitable.

Case 1: $\theta\pi \geq \tau(\tilde{a})$. This also implies that $\theta\tau(a_M^*) \geq \tau(\tilde{a})$. Therefore, in any equilibrium where market beliefs are $a^e = 0$ or $a^e = a_M^*$ if the manager deviates to effort level $\tilde{a}$, she will play the LCSE secondary market optimal outcomes. This is because gains from trade are large enough that it is never optimal ex post to sell a risk-free claim. By construction, the optimal response to LCSE strategies is to exert either effort $a^* = 0$ or $a^* = a_M^* > 0$, depending on initial a^e. Deviation to $\tilde{a}$ is not profitable. As a result, both candidates are equilibria.

Case 2: $\theta\pi < \tau(\tilde{a})$ and $\theta\tau(a_M^*) \geq \tau(\tilde{a})$. In this case, the argument used in the previous case to show existence of the equilibrium with positive effort continues to be true. However, there is a profitable deviation from the equilibrium $\{0, X_H\}$: exert effort $\tilde{a} > 0$ and issue a risk-free claim in secondary markets. This double deviation is possible since $\theta\pi < \tau(\tilde{a})$, which implies that when the manager exerts effort $\tilde{a}$, she will issue a risk-free claim ex post when she holds the good asset. This, in turn, is consistent with $\tilde{a}$, and it is easy to check that it provides the manager higher value than the $\{0, X_H\}$ allocations:

$$V_0(0) = \theta(\pi(X_H - X_L) + X_L) < \max_{a\in[0,1]} a(\tau(a)(X_H - X_L) + \theta X_L)$$
$$+ (1 - a)\theta(\pi(X_H - X_L) + X_L) - C(a) = V_0(\tilde{a}). \tag{E.19}$$

Case 3: $\theta\tau(a_M^*) < \tau(\tilde{a})$. In this scenario, the deviation to effort $\tilde{a}$ is always accompanied by a deviation of the g-type to issue a risk-free claim in secondary markets, since gains from trade are low enough that selling undervalued cash flows is not profitable. In addition, it continues to be true that deviating from $\{0, X_H\}$ is profitable. Thus, it remains to characterize when a deviation from $\{a_M^*, d_M^*\}$ is not profitable, which is the case if $V_0(a_M^*, d_M^*) \geq V_0(\tilde{a}, X_L)$. If this condition does not hold, then there is no equilibrium of the full game.

[Exogenous Quality.] When quality is exogenous, there is only one candidate equilibrium (see Proposition 5). In addition, note that, since secondary market outcomes are independent of the manager's effort choice and of market beliefs, we are always in Case 2: $\theta\bar{\tau} > \bar{\tau}$. Therefore, the candidate is the unique equilibrium and it always exists. In other words, when quality is exogenous, deviating in effort does not affect secondary market outcomes, and for the LCSE strategies, there is a unique effort choice that maximizes ex ante value, and therefore a unique equilibrium.

B. *Complementary Lemmas for Section III*

LEMMA 8: *With D1-Refinements, pooling equilibria in secondary markets do not exist for $a^e > 0$.*

PROOF: Let $a^e > 0$, and suppose that there is a pooling equilibrium in secondary markets for security $F \in \Delta$, and thus $\mu(F) = a$ and $p(F) = \rho(a^e)F(X_H) + (1 - \rho(a^e))F(X_L)$. Then, the g-type manager must be receiving at least the payoff she would obtain by selling the risk-free claim X_L:

$$\theta E_{a^e}[F(X)] - E_a^g[F(X)] \geq (\theta - 1)X_L \tag{E.20}$$

$$(\theta\rho(a^e) - \tau(a))(F(X_H) - F(X_L)) + (\theta - 1)F(X_L) \geq (\theta - 1)X_L. \tag{E.21}$$

Since $F(X_L) \leq X_L$, it must be the case that $\theta\rho(a^e) - \tau(a) \geq 0$. Consider deviation to $F'(X_H) = d$ and $F'(X_L) = F(X_L)$ where d is given by the binding IC of the b-type:

$$(\theta\tau(a^e) - \pi)(d - F(X_L)) = (\theta\rho(a^e) - \pi)(F(X_H) - F(X_L)). \tag{E.22}$$

Note that this implies that $d < F(X_H)$ since $\tau(a^e) > \rho(a^e)$. Thus, the b-type manager is indifferent between deviating and being identified as a g-type or continuing to sell the pooling security F. Finally, if the b-type is indifferent, then the g-type is strictly better off when the deviation is assigned belief $\mu = 1$:

$$\Rightarrow (\theta\tau(a^e) - \tau(a))(d - F(X_L)) > (\theta\rho(a^e) - \tau(a))(F(X_H) - F(X_L)). \tag{E.23}$$

Thus, while the set of beliefs for which the b-type is strictly better off by deviating is empty, the g-type is better off for $(\bar{\mu}, 1]$. By D1-Refinements, this deviation is assigned belief $\mu = 1$. Therefore, it is profitable for the g-type to deviate. Contradiction.

LEMMA 9: *Let $\{a, a^e\}$ be given, and let $\{F_b, F_g\} \in \Delta^2$ be the securities sold in a separating equilibrium in secondary markets. Then F_g is given by the solution to the problem*

$$\begin{aligned} \max_{F\in\Delta} \quad & \theta E_{a^e}^g[F_g(X)] + E_a^g[X - F_g(X)] \\ s.t. \quad & \theta E_{a^e}^g[F_g(X)] + E^b[X - F_g(X)] = u_b^*. \end{aligned} \tag{E.24}$$

where u_b^* is the value of the b-type manager in this equilibrium.

PROOF: Assume not. Then $F_g \notin \{F \in \Delta : F \textit{ solves } \text{(E.24)} \textit{ for } u_b^*\} \equiv \Theta(u_b^*)$. Consider the deviation to security $F'_g \in \Theta(u_b^*)$. The set of beliefs for which the b-type profits from deviating to F'_g is empty, since by the constraint the b-type is indifferent for $\mu = 1$ assigned to the deviation, and is worse-off for $\mu \in [0, 1)$. The g-type, however, strictly prefers selling F'_g for $\mu \in (\underline{\mu}, 1]$, both by construction and by the fact that the original $F_g \notin \Theta(u_b^*)$. By D1, the belief assigned to this deviation is $\mu = 1$, and the g-type deviates. Contradiction.

Appendix F: Understanding the Role of Frictions

To highlight the role of each friction in the model, I solve three problems. First, I consider the model with unobservable screening effort but no private information, that is, the manager does not know the quality of the asset she originates (i.e., there are no manager types). This case highlights the importance of the manager's private information both in implementing the optimal mechanism and in having an equilibrium with positive effort. Second, I consider the case of unobservable effort where the manager's z-type is observed by investors. I refer to this as the model with observable types. Third, I consider the case of observable effort but unobservable z-types, that is, the quality of the originated asset is the manager's private information. I refer to this as the model with observable effort. I show that when effort or z-types are observable, the optimal mechanism is able to implement the first-best allocations, while the market may fail to do so.

A. Hidden Effort Only

In this version of the model, the manager chooses hidden effort a in $t = 0$ to improve the likelihood of financing a good project as in the baseline model. The main difference is that neither the manager nor investors know the quality of the originated asset, that is, there are no manager types. I show that, in this alternative setting, the optimal mechanism implements less ex ante efficiency than when the manager does have private information about her asset quality, and that the equilibrium allocations feature zero screening effort and no retention.

The Optimal Mechanism. When the manager has no private information, there are no z-types in $t = 1$. Thus, the mechanism design problem of Section II is restated as:

$$\max_{a\in[0,1],t\in\mathbb{R}^+,T(\cdot)\in\Delta} \theta t + E_a[T(X)] - C(a) \tag{F.1}$$

$$\text{s.t.} \qquad t \le E_a[X - T(X)] \tag{F.2}$$

$$a = \arg\max_{\hat{a}} \theta t + E_{\hat{a}}[T(X)] - C(\hat{a}). \tag{F.3}$$

Plugging the binding PC of investors and using the first-order approach, we obtain

$$\begin{aligned} &\max_{a,T(\cdot)} \quad \theta E_a[X] - (\theta - 1)E_a[T(X)] - C(a) \\ &\text{s.t.} \quad \rho'(a)(T(X_H) - T(X_L)) = C'(a). \end{aligned} \tag{F.4}$$

It is straightforward that there are no gains from retention in the low-cash-flow realization state (as in the baseline model), and thus $T(X_L) = 0$. Define a new indirect cost: $\tilde{C}_R(a) = (\theta - 1)a\frac{C'(a)}{\rho'(a)}$. Then the effort level implemented in the optimal mechanism is given by the following first-order condition, while retention in the high state is given by the IC for effort constraint:

$$\theta\rho'(a^*)(X_H - X_L) - C'(a^*) = \tilde{C}'_R(a^*) \tag{F.5}$$

$$T(X_H) = \frac{C'(a^*)}{\rho'(a^*)}. \tag{F.6}$$

The main difference with the optimal mechanism effort and retention allocations in the baseline model is given by the new indirect cost, where $\frac{C_R(a)}{\tilde{C}_R(a)} = \tau(a) < 1$. Implementing a given retention level is now more costly, since the manager has to retain in all high-cash-flow states, while before only the g-type retained cash flows. As a result, ex ante efficiency is decreased.

Equilibrium Allocations. In any market equilibrium, market investors correctly infer the manager's effort choice and thus everyone values the asset equally. As a result, since effort is sunk, the manager sells her entire asset to the market to maximize gains from trade in $t = 1$. In the absence of adverse selection to the market, the manager receives price $p_{a^e}(X) = X_L + \rho(a^e)(X_H - X_L)$ for her asset, which is a function of market beliefs and not of the manager's actual effort choice. Thus, the manager's problem in $t = 0$ is:

$$\max_{a\in[0,1]} \theta p_{a^e}(X) - C(a). \tag{F.7}$$

It is easy to see that the manager will always shirk and exert no effort. The only equilibrium is one with zero screening effort and no retention.

B. Hidden-Effort and Observable Types

As in the baseline model, the manager chooses hidden effort a in $t = 0$ given the secondary market outcomes. The manager's z-type is observed by investors. Interestingly, when quality is endogenous, investors' beliefs about the manager's effort choice are important for evaluating the g-type's asset.

The Optimal Mechanism. The problem of the optimal mechanism is as in Section II, but now the IC for type revelation are removed due to observability of z-types.

Consider the following transfers

$$T_g(X) = T_b(X) = 0 \tag{F.8}$$

$$t_g - t_b = \rho'\left(a^*_{FB}\right)(X_H - X_L) \tag{F.9}$$

$$t_b = X_L + \left(\rho\left(a^*_{FB}\right) - a^*_{FB}\rho'\left(a^*_{FB}\right)\right)(X_H - X_L), \tag{F.10}$$

where a^*_{FB} is the first-best effort level. It is easy to verify that these transfers implement first-best. First, there is no retention. Second, by plugging these transfers into the IC for effort, we obtain $\theta(t_g - t_b) - C'(a) = 0$ and thus

$$\theta\rho'\left(a^*_{FB}\right)(X_H - X_L) - C'(a) = 0 \Rightarrow a = a^*_{FB}. \tag{F.11}$$

When types are observable, retention is no longer needed to implement effort. Instead, to generate incentives, $t = 1$ transfers are used. In particular, the difference between $t_b - t_g$ is chosen so that first-best effort levels are implemented. Finally, t_g is chosen to ensure that the PC of investors binds. Therefore, when types are observable, $t = 1$ transfers across manager types are sufficient to implement first-best allocations; costly retention is not needed.

Market Allocations. As in Section III, I solve the problem by backwards induction. First, since in any equilibrium effort is inferred by the market, both manager types sell equity in $t = 1$. As before, however, prices will be a function of market beliefs and not of the direct manager effort choice. Given this, the choice of effort at $t = 0$ is given by

$$a = \arg\max_{\hat{a}} \ \theta \left[\hat{a} E^{g}_{a^e}[X] + (1 - \hat{a}) E^{b}[X]\right] - C(\hat{a}) \tag{F.12}$$

$$\Rightarrow \ \theta \left(\tau(a) - \pi\right)(X_H - X_L) - C'(a) = 0, \tag{F.13}$$

where the equilibrium condition $a^e = a$ is imposed after taking first-order conditions. When quality is exogenous, $\rho'(a) = \bar{\tau} - \pi$ and thus the market implements first-best allocations. This is because when prices do not depend on effort, the z-type manager is fully compensated in secondary markets for her effort choice and there is neither an effort nor a retention externality.

When quality is endogenous, however, the market always under-exerts effort relative to first-best. The difference $\tau(a) - \pi < \rho'(a) = \tau(a) - \pi + a\tau'(a)$, captures the effort externality that arises in this scenario since effort is not observable. The retention externality is not present since there is no adverse selection when types are observable.

C. Observable Effort and Unobservable Types

In this section, I assume that the screening effort a is observed by market investors but that the manager's z-type is the manager's private information.

Optimal Mechanism. Since effort is observable, the problem of the optimal mechanism is as in Section II, but without the IC for effort. Consider the following transfers and effort choice (which can be directly implemented since it is observable):

$$T_g(X) = T_b(X) = 0 \tag{F.14}$$

$$t_g = t_b = X_L + \rho\left(a^*_{FB}\right)(X_H - X_L) \tag{F.15}$$

$$a^* = a^*_{FB}. \tag{F.16}$$

It is easy to check that the previous transfers implement the first-best allocations and that they satisfy all the optimal mechanism constraints. Under the proposed mechanism, there is pooling of types in the market for equity, since they both receive the same $t = 1$ transfers. The first-best allocation for effort is implemented directly since effort is observed, and transfers can be made contingent on the observation of the first-best effort level. Therefore retention of cash flows is not required.

Market Allocations. When effort is observable but manager z-type is not, the LCSE is the unique equilibrium in secondary markets, where the debt level issued by the g-type is

$$d(a) = X_L + \frac{(\theta - 1)\pi}{\theta\tau(a) - \pi}(X_H - X_L). \tag{F.17}$$

In contrast to the baseline model, when effort is observable, secondary market prices and debt levels directly depend on the effort choice of the manager; there is no effort externality. Since the manager understands this, her ex ante effort choice is given by

$$a_M^* = \arg\max_a \theta E_a[X] - (\theta - 1)a\tau(a)(X_H - d(a)) - C(a). \tag{F.18}$$

The first-order condition that characterizes the manager's effort choice is given by

$$\theta\rho'(a)(X_H - X_L) - C'(a) - \underbrace{[(\theta - 1)(a\tau'(a) + \tau(a))(X_H - d(a)) - \theta a\tau(a)d'(a)]}_{\text{Distortion Relative to First-Best}\equiv\mathcal{D}(a)} = 0. \tag{F.19}$$

In contrast to the baseline case, the presence of adverse selection in secondary markets reduces incentives to exert effort when the latter is observable. From equation (F.17) we can see that $d'(a) \leq 0$, which implies that $\mathcal{D}(a) > 0$ for all $a \in (0, 1)$: the market under-exerts effort relative to the optimal mechanism, which implements first-best. The effort choice is distorted because when types are unobservable, the g-type manager has to retain cash flows to signal her quality to investors. The presence of this retention implies that (1) the return on effort is reduced, since gains from trade are forfeited for the cash flows that are retained, and (2) in the endogenous-quality case, the manager further distorts her effort choice to affect how much debt is issued in secondary markets.

REFERENCES

Agarwal, Sumit, Yan Chang, and Abdullah Yavas, 2012, Adverse selection in mortgage securitization, *Journal of Financial Economics* 105, 640–660.

Aghion, Philippe, Patrick Bolton, and Jean Tirole, 2004, Exit options in corporate finance: Liquidity versus incentives, *Review of Finance* 8, 327–353.

Ashcraft, Adam B., Kunal Gooriah, and Amir Kermani, 2014, Does skin-in-the-game affect security performance?, Working paper. Available at SSRN 2437574.

Berndt, Antje, and Anurag Gupta, 2008, Moral hazard and adverse selection in the originate-to-distribute model of bank credit, *Journal of Monetary Economics* 56, 725–743.

Bhide, Amar, 1993, The hidden costs of stock market liquidity, *Journal of Financial Economics* 34, 31–51.

Biais, Bruno, and Thomas Mariotti, 2005, Strategic liquidity supply and security design, *Review of Economic Studies* 72, 615–649.

Botsch, Matthew J., and Victoria Magdalena Vanasco, 2016, Learning by lending, Working paper. Available at SSRN 2619961.

Calem, Paul, Christopher Henderson, and Jonathan Liles, 2011, Cherry picking in subprime mortgage securitizations: Which subprime mortgage loans were sold by depository institutions prior to the crisis of 2007?, *Journal of Housing Economics* 20, 120–140.

Chari, V.V., Ali Shourideh, and Ariel Zetlin-Jones, 2014, Reputation and persistence of adverse selection in secondary loan markets, *American Economic Review* 104, 4027–4070.

Chemla, Gilles, and Christopher A. Hennessy, 2014, Skin in the game and moral hazard, *Journal of Finance* 69, 1597–1641.

Cho, In-Koo, and David M. Kreps, 1987, Signaling games and stable equilibria, *Quarterly Journal of Economics* 102, 179–221.

Cho, In-Koo, and Joel Sobel, 1990, Strategic stability and uniqueness in signaling games, *Journal of Economic Theory* 50, 381–413.

Chodorow-Reich, Gabriel, 2014, The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis, *Quarterly Journal of Economics* 129, 1–59.

Coffee, John, 1991, Liquidity versus control: The institutional investor as corporate monitor, *Columbia Law Review* 91, 1277–1368.

Dell'Ariccia, Giovanni, Deniz, Igan and Luc Laeven, 2012, Credit booms and lending standards: Evidence from the subprime mortgage market, *Journal of Money, Credit and Banking* 44, 367–384.

DeMarzo, Peter M., 2005, The pooling and tranching of securities: A model of informed intermediation, *Review of Financial Studies* 18, 1–35.

DeMarzo, Peter M., and Darrell Duffie, 1999, A liquidity-based model of security design, *Econometrica* 67, 65–99.

Dewatripont, Mathias, and Jean Tirole, 1994, A theory of debt and equity: Diversity of securities and manager-shareholder congruence, *Quarterly Journal of Economics* 109, 1027–1054.

Downing, Chris, Dwight Jaffee, and Nancy Wallace, 2008, Is the market for mortgage-backed securities a market for lemons?, *Review of Financial Studies* 22, 2457–2494.

Elul, Ronel, 2016, Securitization and mortgage default, *Journal of Financial Services Research* 49, 281–309.

Fama, Eugene F., 1985, What's different about banks?, *Journal of Monetary Economics* 15, 29–39.

Faure-Grimaud, Antoine, and Denis Gromb, 2004, Public trading and private incentives, *Review of Financial Studies* 17, 985–1014.

Fender, Ingo, and Janet Mitchell, 2009, The future of securitisation: How to align incentives?, *BIS Quarterly Review*. Available at SSRN: https://ssrn.com/abstract=1472970.

Financial Stability Board (FSB), 2010, Overview of progress in the implementation of the G20 recommendations for strengthening financial stability, Report to G20 Leaders. Available at www.financialstabilityboard.org/publications/r_100627c.pdf.

Gorton, Gary B., and George G. Pennacchi, 1995, Banks and loan sales. Marketing nonmarketable assets, *Journal of Monetary Economics* 35, 389–411.

Hartman-Glaser, Barney, 2017, Reputation and signaling in asset sales, *Journal of Financial Economics* 00, 1–21.

Hartman-Glaser, Barney, Tomasz Piskorski, and Alexei Tchistyi, 2012, Optimal securitization with moral hazard, *Journal of Financial Economics* 104, 186–202.

Hébert, Benjamin, 2015, Moral hazard and the optimality of debt. Available at SSRN: https://ssrn.com/abstract=2185610 or https://doi.doi.org/10.2139/ssrn.2185610.

Holmström, Bengt, and Jean Tirole, 2011, *Inside and Outside Liquidity* (MIT Press, Cambridge, MA).

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

Jaffee, Dwight, Anthony Lynch, Matthew Richardson, and Stijn Van Nieuwerburgh, 2009, Mortgage origination and securitization in the financial crisis, in Viral Acharya, and Matthew Richardson, eds.: *Restoring Financial Stability*, chapter 1. Hoboken, New Jersey.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil, 2014, Liar's loan? Effects of origination channel and information falsification on mortgage delinquency, *Review of Economics and Statistics* 96, 1–18.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vig Vikrant, 2010, Did securitization lead to lax screening? Evidence from Subprime Loans, *Quarterly Journal of Economics* 125, 307–362.

Leland, Hayne E., and David H. Pyle, 1977, Informational asymmetries, financial structure, and financial intermediation, *Journal of Finance* 32, 371–387.

Malherbe, Frédéric, 2012, Market discipline and securitization, Working paper, London Business School.

Maug, Ernst, 1998, Large shareholders as monitors: Is there a trade-off between liquidity and control?, *Journal of Finance* 53, 65–98.

Medrano, Luis Angel, and Xavier Vives, 2004, Regulating insider trading when investment matters, *Review of Finance* 8, 199–277.

Mian, Atif, and Amir Sufi, 2009, The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis, *Quaterly Journal of Economics* 124, 1449–1496.

Mikkelson, Wayne H., and M. Megan Partch, 1986, Valuation effects of security offerings and the issuance process, *Journal of Financial Economics* 15, 31–60.

Mondragon, John, 2015, Household credit and employment in the great recession, Working paper. Available at SSRN 2521177.

Myers, Stewart C., and Nicholas S. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* 13, 187–221.

Nachman, David, and Thomas Noe, 1994, Optimal design of securities under asymmetric information, *Review of Financial Studies* 7, 1–44.

Parlour, Christine A., and Guillaume Plantin, 2008, Loan sales and relationship banking, *Journal of Finance* 63, 1291–1314.

Plantin, Guillaume, 2009, Learning by holding and liquidity, *Review of Economic Studies* 76, 395–412.

Plantin, Guillaume, 2011, Good securitization, bad securitization, Institute for Monetary and Economic Studies, Bank of Japan.

Qian, Jun Q.J., Philip E. Strahan, and Zhishu Yang, 2015, The impact of incentives and communication costs on information production and use: Evidence from bank lending, *Journal of Finance* 70, 1457–1493.

Shleifer, Andrei, and Robert W. Vishny, 1997, A survey of corporate governance, *Journal of Finance* 52, 737–783.

Slovin, Myron B., Marie E. Sushka, and John A. Polonchek, 1999, An analysis of contagion and competitive effects at commercial banks, *Journal of Financial Economics* 54, 197–225.

Yang, Ming, and Yao Zeng, 2014, Security design in a production economy with flexible information acquisition, Working paper. Available at SSRN 2194194.