Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Market Freeze and Bank Capital Structure Heterogeneity

Fenghua Song, Anjan V. Thakor

To cite this article:
Fenghua Song, Anjan V. Thakor (2022) Market Freeze and Bank Capital Structure Heterogeneity. Management Science

Published online in Articles in Advance 24 Mar 2022

. https://doi.org/10.1287/mnsc.2022.4374

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–21
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Market Freeze and Bank Capital Structure Heterogeneity

Fenghua Song,[a,*] Anjan V. Thakor[b,c]

[a] Smeal College of Business, Pennsylvania State University, University Park, Pennsylvania 16802; [b] Olin Business School, Washington University in St. Louis, St. Louis, Missouri 63130; [c] European Corporate Governance Institute (ECGI), 1000 Brussels, Belgium
*Corresponding author

Contact: song@psu.edu, https://orcid.org/0000-0002-1043-1248 (FS); thakor@wustl.edu, https://orcid.org/0000-0002-3919-6456 (AVT)

Received: December 8, 2021
Revised: February 17, 2022
Accepted: February 19, 2022
Published Online in Articles in Advance: March 24, 2022

https://doi.org/10.1287/mnsc.2022.4374

Abstract. We develop a theory wherein a priori identical banks may trade loans in a search market with reusable information. The equilibrium is unique, but its nature depends on the probability of a future market state. When the probability of a boom is high, all banks hold no equity and do no screening. When this probability is low, all banks choose a high level of equity and screen loans. For intermediate probability values, the equilibrium is heterogeneous, with some banks posting equity and screening and others avoiding equity and screening. This endogenously arising heterogeneity generates interbank trading. The credit market is partially frozen in a recession: only high-capital banks have continued funding access. Low-capital banks obtain funding by selling legacy loans to banks with "financial muscle," so market funding is reallocated from low-capital to high-capital banks.

History: Accepted by Bruno Biais, finance.

Keywords: market freeze • bank capital • information reusability • short-term funding reallocation

1. Introduction

During crises, many financial institutions get frozen out of short-term funding markets. If such funding dry-up is systemic, it can lead to fire sales with downward price spirals (Shleifer and Vishny 1992, Allen and Gale 1994) and necessitate government intervention (Philippon and Skreta 2012, Tirole 2012). In models designed to characterize these phenomena, *all* agents trade less because of adverse selection. However, recent evidence shows that the 2007–2009 crisis was different. Although some institutions were unable to access funding markets, others actually *expanded* access. Specifically, short-term funding markets reallocated liquidity from weak to strong banks. Using transaction-level data on short-term unsecured certificates of deposit in European markets, Pérignon et al. (2018) document that, although many banks experienced sudden funding dry-ups during 2008–2014, there was no marketwide freeze: Banks with higher capital actually *increased* their short-term uninsured funding, whereas those with lower capital reduced funding. Similarly, Boyson et al. (2014) find that U.S. commercial bank funding did not dry up during 2007–2009, but the market forced weak (low-capital) banks to borrow less. Berger and Bouwman (2013) show that during crises, high-capital banks grow their loan and deposit market shares, whereas low-capital banks shrink, and one channel through which this occurs is acquisition of the assets of low-capital banks by high-capital banks. That is, high-capital banks appear to enjoy significant competitive advantages during crises that stem from reallocations because of trading among banks that are heterogeneous in asset qualities and capital ratios. What accounts for these patterns in the data?

We build a theory consistent with these findings. In a two-period model, *ex ante identical* banks make loans in the first period. Loans have a priori unknown credit risks that banks, as screening specialists, can discover at a cost. However, banks will not invest in screening without sufficient equity, as in Holmstrom and Tirole (1997). Banks' capital structure and screening choices at the start of the first period are endogenous.

After loans are made, a macroeconomic state, boom or recession, is realized and publicly observed at an interim date. The quality of a legacy loan made earlier is higher in a boom than in a recession. In the second period, banks raise financing at that interim date for new (positive net present value (NPV)) projects. As in Tirole (2012), limited pledgeability prevents raising financing by solely issuing claims against the new projects, so part of the claims to finance these projects must be issued against legacy assets (first-period loans). However, investors recognize that banks with insufficient capital did not screen their legacy loans, so they are unwilling to provide financing for new projects to such banks if the realization of the macroeconomic state is a recession. Investors also know that high-capital banks invested in screening and hence have better legacy assets on average than low-capital banks. Thus, unlike low-capital banks, high-capital banks are able to raise financing for their new projects even in a recession. No bank experiences a funding freeze in a boom because the qualities of all banks' legacy loans are sufficiently high. However, in a recession, low-capital banks experience a freeze: Their low-quality legacy assets are unable to support financing for new projects. By contrast, high-capital banks enjoy continued funding access because their legacy assets are of sufficiently high quality even in downturn. Thus, whether a funding freeze is experienced (by some banks) is linked to banks' earlier endogenous screening decisions and their chosen capital levels.

We first examine a benchmark case in which interbank trading of (legacy) loans is prohibited. The resulting equilibrium is unique and has all (a priori identical) banks behaving identically: either all banks hold equity and screen or no bank raises equity and screens. Whether we obtain a screening or no-screening benchmark equilibrium depends on the probability of a boom, θ, a deep parameter of the model.

We then introduce interbank trading in legacy loans in an over-the-counter market involving search and matching. Now, for certain values of θ, a heterogeneous equilibrium emerges in which some banks hold positive capital and screen, whereas others post no capital and avoid screening, despite all banks being identical a priori. High-capital banks act as "expert buyers" who can evaluate the quality of other banks' legacy assets. Because the screening investment made by these high-capital banks in the first period is (at least partially) reusable across loans with similar characteristics,[1] each can use its existing screening capability to further evaluate and purchase good legacy loans from those no-capital banks. Such trading provides the (otherwise frozen) no-capital banks funds to invest in their new assets in the second period, so they are no longer marooned by their own illiquidity.

This result differs sharply from previous theories dealing with markets frozen by adverse selection. In Philippon and Skreta (2012) and Tirole (2012), the government helps unfreeze the credit market by buying the *worst* assets. This increases the average quality of assets remaining in the market, helping revive market financing. By contrast, in our model expert buyers (high-capital banks) directly buy the *best* assets (identified by their screening) and fund the purchases by selling claims against these assets (implicitly) "certified" as being high quality by virtue of the buyers' willingness to purchase them.

The model has the following additional findings. First, when θ (boom probability) is sufficiently high or low, the equilibrium is unique but remains homogeneous (as in the benchmark no-trading equilibrium) despite the availability of interbank trading. For intermediate θ values, a unique heterogeneous equilibrium arises to exploit gains to trade that heterogeneity generates, and the range of these values expands with higher trading efficiency. Second, in the heterogeneous equilibrium the masses of high-capital banks and no-capital banks (hence, interbank market tightness) are endogenously determined. We call this a "general equilibrium" because there is nothing to distinguish banks ex ante, and their capital structure and asset quality heterogeneities both arise endogenously because of the anticipation of trading. This is in contrast to partial equilibrium models wherein bank heterogeneity is taken as exogenously given and then its implications are analyzed (Goldstein et al. 2020). Third, trading provides a loan buyer (high-capital bank) an extra incentive to invest in screening to better identify the seller's loan quality. This lessens the reliance on costly equity as "skin in the game" to incentivize screening and is socially valuable. We show that this incentive role of trading is stronger with a lower buyer/seller ratio in the interbank market. However, no individual bank internalizes the social value of this lower ratio when making its own decision to become a buyer or seller. As a result, the heterogeneous equilibrium involves an excessive mass of high-capital banks (buyers), each holding too much equity, relative to the social optimum. This inefficiency is not because of the standard congestion externality; we say more in Section 3.3.

Heterogeneity among a priori identical banks arises *endogenously* as an equilibrium phenomenon because heterogeneity is essential for trading, and trading benefits banks.[2] This happens because trading enables no-capital banks to avoid costly equity and screening and high-capital banks to profit from loan purchases during times of stress. Essentially, we can think of no-capital banks "outsourcing" screening to high-capital banks and then seeking financing from high-capital banks rather than directly from the market that is frozen for them in a recession. Because high-capital banks raise their financing from the market, they act as de facto intermediaries between no-capital banks and the market, enabling endogenous liquidity provision to no-capital banks. This resembles "global value chains" wherein different countries engage in different parts of the production process and then trade, thereby enhancing welfare by optimizing the use of resources. In our model, the resources that are optimized are costly equity and screening.

There is no economic rationale for regulatory capital requirements in our model; a further discussion of this is in Section 4.1. Our results nonetheless raise the obvious question about what effect capital requirements would have in our model. Given a vast literature providing a host of reasons for capital requirements (see Merton (1977) and the subsequent literature on how capital requirements can be welfare enhancing), in a model extension, we take a minimum capital requirement as given to examine its effect. We find that even a capital requirement that is binding only on selling banks and not on buying banks can induce *both* types of banks to increase their capital and a sufficiently high capital requirement may eliminate the heterogeneous equilibrium. Thus, the safety net and other benefits of capital requirements need to be weighed against the reduction in welfare because of the elimination of the heterogeneous equilibrium.

Our result that heterogeneity can emerge in equilibrium among a priori identical banks has testable predictions. Specifically, heterogeneity is more likely with higher interbank market trading efficiency and with higher cross-sectional information reusability (i.e., among banks lending to more similar firms). These predictions can be tested to assess the empirical merit of our explanation for heterogeneity compared with other ways (e.g., regulatory restrictions) in which heterogeneity may arise.

Our analysis also has policy implications. First, heterogeneity is important for banking stability (Goldstein et al. 2020), so a deeper understanding of the economic forces that facilitate such heterogeneity is valuable for macroprudential regulators. Second, besides its stability benefits, we show that heterogeneity also has trading benefits and helps economize on costly equity and screening. These benefits run counter to the general direction of global bank regulation that seeks to homogenize banks and not create heterogeneity. Therefore, the analysis has important regulatory implications. Third, heterogeneity can be created in many ways, including with regulatory intervention. For example, Acharya and Yorulmazer (2008) show that liquidity assistance to some banks during a crisis can provide ex ante incentives for banks to differentiate. What we show is that there are *natural equilibrium forces* that generate heterogeneity without a regulatory assist (or other exogenous forces), and those natural forces are particularly strong with an efficient interbank market for trading. In those instances, regulatory initiatives that impede these forces can be especially distorting in ways not earlier recognized.

Our analysis rests on several features ubiquitous in banking theories. First, savers invest indirectly in borrowers through banks, and banks specialize in screening credit risks (Ramakrishnan and Thakor 1984, Boyd and Prescott 1986, Coval and Thakor 2005). Second, screening investments display information reusability, which is at the heart of numerous banking theories (Chan et al. 1986, Sharpe 1990, Novo-Peteiro 2000) and plays a central role in relationship banking (Boot 2000). Third, we examine capital requirements, a regulation unique to banking. Fourth, bank assets can be traded relatively free of frictions other than adverse selection, and this trading occurs in a recession or crisis (in the heterogeneous equilibrium). The ease of trading is because bank assets are financial claims that require owners to only collect cash flows from these assets but not actually operate them. By contrast, real-sector firms hold hard assets whose value depends on whether the owner has expertise in operating these assets, which causes the assets to *not* be frequently traded in liquid secondary markets. There is ample evidence of banks using the interbank market for loan sales extensively to cope with negative funding shocks. Irani and Meisenzahl (2017) find, consistent with our prediction, that banks purchasing loans during crises are high-capital banks. Irani et al. (2021) document that shocks to capital requirements cause low-capital banks to shed loans, whereas Gambacorta and Shin (2018) show that higher capital enables banks to increase loan growth, consistent with the earlier-cited Berger and Bouwman (2013) evidence that high-capital banks acquire assets from low-capital banks during crises and the evidence in Granja et al. (2017) on the resolution of bank failures by the Federal Deposit Insurance Corporation (FDIC). Thus, our prediction of loan sales from low-capital to high-capital banks during crises is consistent with the empirical evidence, and it seems to be unique to banking. For example, such transactions did not occur among automobile companies during the 2007–2009 crisis, although they were just as stressed as banks.

This paper is related to the fire-sales literature, pioneered by Allen and Gale (1994) and Shleifer and Vishny (1992, 2011). In these models, because of cash-in-the-market pricing constraints or the necessity of selling to nonexpert second-best users, asset prices fall with sales.[3] By contrast, in our model, banks with financial muscle are well-capitalized specialist buyers, so asset sales do not lead to an inefficiency per se.

Also related is the literature on government intervention in markets frozen by adverse selection, both in static (Philippon and Skreta 2012, Tirole 2012, Jorge and Kahn 2017) and dynamic settings (Camargo and Lester 2014, Chiu and Koeppl 2016). Our paper differs in two important aspects. First, we introduce informed buyers of legacy assets who provide an alternative to both direct market financing by asset sellers and government intervention. Second, we endogenize bank equity choice. These features lead to a novel interaction between bank capital structure and illiquidity risk.

Another relevant research strand is on bank failures and runs which shows that heterogeneity enhances financial stability (Choi 2014, Goldstein et al. 2020). What our paper shares with them is that bank heterogeneity is beneficial. A key difference is that they take heterogeneity as exogenously given, whereas in our model it arises endogenously, so we can link the origin of heterogeneity to deep model parameters (boom probability and interbank market efficiency). Also, our focus is not on runs, but how a credit market freeze can be alleviated *without* government intervention because of trading made possible by capital structure heterogeneity among a priori identical banks.

The bank capital literature (Holmstrom and Tirole 1997; Allen et al. 2011, 2015; Mehran and Thakor 2011; Carletti et al. 2020) is also related. Unlike us, this literature does not examine how bank capital affects future liquidity crises. In that sense, our paper also complements recent empirical research documenting the beneficial effects of higher bank capital in conducting monetary policy (Gambacorta and Shin 2018), enhancing bank survival probability and market share (Berger and Bouwman 2013), and reducing systemic risk (Bostandzic and Weiß 2018).

Finally, the literature on liquidity hoarding as a precautionary response to possible future funding stress is also related (Acharya and Skeie 2011, Acharya et al. 2012, Gale and Yorulmazer 2013). In contrast to this literature, endogenously arising capital heterogeneity and the consequent interbank trading in our model replace central bank intervention as unfreezing mechanisms. Moreover, although the earlier literature focuses on the ex ante misallocation of resources because of loanable funds being diverted to cash (Acharya et al. 2013), our analysis has no such distortion because of the high-capital bank holding more equity, and heterogeneity always elevates welfare.

Section 2 describes the model. Section 3 has the analysis (proofs in Appendix A). Section 4 examines robustness and extension (technical details in Appendix B) and discusses implications. Section 5 concludes.

2. Model

Agents are risk neutral, and the risk-free rate of interest is zero. There are three dates: $t=0$, 1, 2. There are numerous atomistic and a priori identical banks, each making three decisions at $t=0$: (i) capital structure (i.e., the mix of deposits and equity); (ii) whether to invest in a technology to screen a loan's quality; and (iii) whether to approve or reject the loan based on (ii).[4] Although (i) and (iii) are publicly observable, (ii) is private to the bank.

2.1. Loans

Loans can be one of two types: good (G) or medium (M); each has financing need of I at $t=0$ and pays off only at $t=2$. A type-G loan always returns $X > I$. An M loan's payoff depends on the realization of a macroeconomic state, $s \in \{b, r\}$, at $t=1$, where b stands for boom and r stands for recession. The M loan's date 2 payoff is X if $s=b$, but zero if $s=r$. It is common knowledge that $\Pr(s=b)=\theta \in [0, \overline{\theta}]$, with $\overline{\theta} < \frac{I}{X}$. Therefore, $\theta X < I$, so an M loan is not creditworthy at $t=0$. Loan type is initially unknown to anyone at $t=0$. The common prior belief is that $\Pr(G)=\Pr(M)=\frac{1}{2}$.[5]

2.2. Screening

The bank can screen a loan to determine its quality. By incurring a cost c, screening at $t=0$ perfectly reveals the loan's type, G or M.[6] The investment in screening and the outcome of screening are private to the bank. Conditional on screening, the bank lends at $t=0$ only when screening reveals the loan is G.[7] The date 0 screening investment c further allows the bank to screen another loan at $t=1$ with perfect precision at an additional but reduced cost κc, with $\kappa \in [0,1)$. A bank that extends a loan without screening at $t=0$ does not observe the loan's type even at $t=1$. Screening is not contractible.

The assumption $\kappa < 1$ reflects cross-sectional information reusability[8]: A bank lending to a borrower in one industry acquires industry-specific information that is not only relevant for that borrower but also for other borrowers in the same or related industries. Chan et al. (1986) introduced the idea of information reusability in bank lending and focused on both intertemporal and cross-sectional information reusability, arguing that such information reusability is an essential part of the role banks play as information processors. Although intertemporal information reusability, which gives the bank the ability to use at date $t+1$ proprietary information about the borrower gleaned at date t, has played a central role in relationship banking,[9] the literature on bank specialization has focused on the role of cross-sectional information reusability (Paravisini et al. 2015, Greenbaum et al. 2019, Di and Pattison 2020). There is considerable empirical evidence showing that information reusability benefits borrowers via lower loan spreads (Bharath et al. 2011) and increases profits of relationship lenders (Bharath et al. 2007b).[10]

Our analysis relies on information reusability. Although information reusability may be pertinent for nonbanks as well, it has not been assigned a pivotal role in theories of firms like it has been in banking. This may be because the existence of firms does not crucially depend on information reusability in (most) theories of firms, whereas many banking theories rely on it (Coval and Thakor 2005, Bolton et al. 2016b).[11]

2.3. Bank Capital Structure

The bank raises deposits $D \in [0, I]$ at $t=0$ (legacy debt) and may again at $t=1$ (new debt). Both the legacy debt and new debt mature and are repaid at $t=2$. The legacy debt is more senior than the new debt. The deposit market is perfectly competitive, so repayment obligations on the legacy and new debt are set to yield depositors a zero expected return. Default on debt at $t=2$ causes the bank to lose its charter value, $Y(\theta)$, which (as viewed at $t=0$) is a function of the probability, θ, of the date 1 macro state being a boom. We say more about $Y(\theta)$ later.

Deposits are uninsured but cheaper than equity. This standard assumption can be justified in many ways, for example, taxes, and larger informational frictions for equity (Myers and Majluf 1984). Specifically, although deposits are priced so that the expected debt repayment equals the amount of deposits raised, there is a cost, ψE,[12] to bank insiders for providing E in equity. Banks maximize the wealth of insiders.

Each bank's capital structure is publicly observable. All banks are a priori identical in our model. We will show, however, that when interbank loan trading at $t = 1$ is allowed, there are circumstances in which banks choose different capital structures at $t = 0$.

2.4. New Asset

The macro state s is publicly realized at $t = 1$. Each bank may raise an additional I at $t = 1$ to fund a new asset that also yields $X > I$ (with certainty) at $t = 2$. As in Tirole (2012), the pledgeability of this new asset's cash flow is so limited that its financing I cannot be raised if only this new asset is available to repay financiers. Specifically, the new asset's pledgeable cash flow is δI, with $\delta \in (0,1)$, so the remaining financing $(1-\delta)I$ is made possible by giving date 1 financiers access to the legacy loan's date 2 cash flow. By contrast, the legacy loan's payoff is fully pledgeable.

As an alternative to raising market financing for the new asset by issuing new debt (partly) against the legacy loan's date 2 cash flow, a bank may sell its legacy loan to another bank. The bank that purchases the loan can avail its date 0 investment in the screening technology to inspect the loan. Because of the cross-sectional reusability of the screening technology, the buyer can identify the seller's loan type (G or M) at $t = 1$ at a lower cost than it incurred in acquiring the technology at $t = 0$.

2.5. Loan Trading

Loan trading occurs at $t = 1$, after the macro state, $s \in \{b, r\}$, is realized. Banks interact in an over-the-counter market for legacy loans. We model loan trading with a (one-shot) random search model. The total mass of banks, N, is fixed, among which the mass of buyers is N_b, and the mass of sellers is N_s; N_b and N_s will be determined in the analysis. Let the buyer/seller ratio, $n = \frac{N_b}{N_s}$, denote the market tightness, which determines the probability that buyers and sellers are matched. The probability a seller meets a buyer is $\alpha(n)$, an increasing and concave function with $\alpha(0) = 0$. The probability a buyer meets a seller is $\frac{\alpha(n)N_s}{N_b} = \frac{\alpha(n)}{n}$, which is decreasing in n.

As shown later, trading may occur only in a recession ($s = r$). When $s = r$, after a buyer and a seller are matched, the buyer inspects the seller's loan. If the buyer verifies the seller's loan is G, they negotiate the price p, which is determined according to Nash bargaining with β being the buyer's bargaining power, and $1-\beta$ the seller's bargaining power. We assume that the well-known Hosios condition is satisfied (Hosios 1990), which states that a bank's bargaining power is commensurate with its contribution to matching. Specifically, a buyer's bargaining power equals the elasticity of $\alpha(n)$ with respect to the market tightness n, that is, $\beta = \frac{n\alpha'(n)}{\alpha(n)}$. If the inspection reveals the seller's loan is M, whose date 2 payoff will be zero in the recession, the buyer walks away without a purchase.[13]

The buyer finances its loan purchase by issuing new debt to depositors at $t = 1$, pledging the purchased loan's date 2 cash flow as repayment. The new depositors can be sure that the buyer has verified the purchased loan is G; no buyer would buy a loan identified to be M.[14] The seller uses its sale proceeds, p, to provide $(1-\delta)I$ of the financing it needs for the new asset, with the remaining financing, δI, coming from new debt supported by pledging the new asset's date 2 cash flow. At $t = 2$, both the buyer and the seller first repay their legacy debt (raised at $t = 0$), and then the new debt.

2.6. Model Summary and Parametric Assumptions

Table 1 summarizes the model's timeline.

Assumption 1. *We make the following assumption on bank debt capacity:*

$$(2-\delta)I < X < \min\{2I, 2[(1+\overline{\theta})^{-1} + (1-\delta)]I\}. \quad (1)$$

This condition helps us focus on cases of interest. The restriction $(2-\delta)I < X$ means that if financiers knew with certainty that the loan was type G, then the total financing needed for the legacy loan, I, and the new asset, $(1-\delta)I$,[15] could be raised upfront at $t = 0$ against the G legacy loan's pledgeable date 2 cash flow X. The assumption $X < 2I$ ensures the date 0 NPV of an unscreened loan, $[\theta + 0.5(1-\theta)]X - I$, is negative if the boom probability θ is sufficiently low. With $X < 2[(1+\overline{\theta})^{-1} + (1-\delta)]I$, an unscreened legacy loan can never be expected to enable repayment of the legacy debt and also fund the new asset in a recession even when the boom probability is at its highest value ($\theta = \overline{\theta}$). Note $\theta \in [0, \overline{\theta}]$, with $\overline{\theta} < \frac{I}{X}$. From (1), we have $\frac{1}{2} < \frac{I}{X} < \frac{1}{2-\delta}$. Thus, the analysis can be conducted for a wide range of $\theta \in [0, \overline{\theta}]$ with $\overline{\theta} > \frac{1}{2}$, capturing various scenarios wherein a boom can be either more or less probable than a recession.

Assumption 2. *The bank's charter value $Y(\theta)$ is increasing and concave in θ, satisfying*

$$Y'(\overline{\theta}) > \frac{2(X - I + c)}{(1-\overline{\theta})^2}, \quad (2)$$

$$Y(0) < 2c. \quad (3)$$

Table 1. Sequence of Events

Date	Events
$t = 0$	• Each bank chooses its publicly observable capital structure. • Each bank privately invests in screening, and then makes publicly observable loan approval/rejection decision.
$t = 1$	• Macro state s is realized and becomes public. This determines the payoff of the M loan. • Banks trade legacy loans. Purchasing banks can use their date 0 screening investment to screen loans they purchase at a lower cost because of information reusability. • If possible, each bank raises financing for a new asset to add to its loan portfolio.
$t = 2$	• Both the legacy loan and new asset (if funded at $t = 1$) pay off. • The legacy debt and new debt are paid off. Bank insiders collect the rest.

As shown later, (2) and $Y''(\theta) < 0$ imply that the required equity to induce bank screening is decreasing in θ (boom probability), and (3) ensures that some equity is necessary (as "skin in the game") in addition to bank charter value to provide the bank with the incentive to screen. The specification that a higher expectation of a boom raises bank charter value ($Y'(\theta) > 0$) resonates with the idea that charter values reflect expected future profitability, which is higher when the economy is doing better (Begenau et al. 2020). There is also ample empirical evidence that bank charter values are indeed higher when economic conditions are more favorable (Saunders and Wilson 2001, Furlong and Kwan 2006). As for (3), the point that equity is needed to provide banks with screening and monitoring incentives has been widely used in banking theories (Holmstrom and Tirole 1997, Coval and Thakor 2005) with strong empirical support (Bouwman 2019, Bhat and Desai 2020).

Although the core intuition underlying the model's mechanism does not rely on bank charter value (it mostly reinforces the positive effect of equity in strengthening screening incentives), modeling the charter value as a smooth function of θ (with Assumption 2), $Y(\theta)$, allows us to focus on the economically meaningful and practically relevant case in which the equilibrium equity is decreasing in θ.[16]

Assumption 3. *Screening is economically valuable*:

$$X - I - c - \psi[2c - Y(0)] > 0. \tag{4}$$

As will become clear later, (4) ensures that the bank's profit from screening, net of the screening cost and the cost of equity needed to incentivize screening, is positive for any $\theta \in [0, \overline{\theta}]$.

Assumption 4. *The value-added of screening diminishes over boom prospects:*

$$\psi\left[Y'(0) - \frac{2(X - I + c)}{(1 - \overline{\theta})^2}\right] + 0.5Y'(0) \leq X + 0.5[Y(0) + I]. \tag{5}$$

This is a sufficiency condition for the value-added of bank screening relative to no-screening to decline with the boom probability θ. Such monotonicity allows us to characterize the equilibrium neatly.[17]

Assumption 5. *For tractability, we follow the monetary economics literature (Kiyotaki and Wright 1993) and use a common specification of the meeting technology*

$$\alpha(n) = \frac{\lambda n}{1 + n} \tag{6}$$

to parameterize the probability a seller meets a buyer in the interbank market for legacy loans. The efficiency of that market is captured by $\lambda \in [0, 1]$, with a larger λ corresponding to higher efficiency. Therefore, the probability a buyer is matched with a seller is $\frac{\alpha(n)}{n} = \frac{\lambda}{1+n}$, and the buyer's bargaining power is $\beta = \frac{n\alpha'(n)}{\alpha(n)} = \frac{1}{1+n}$, whereas the seller's bargaining power is $1 - \beta = \frac{n}{1+n}$.

3. Results

The analysis in this section assumes $\kappa = 0$ (information is fully reusable) to convey the essence of our model in the sharpest way. Appendix B.1 shows that results are qualitatively unchanged as long as κ is not too large (there is nontrivial information reusability). In our base model, banks are unregulated with no capital requirements. In an extension in Appendix B.2, we examine the impact of capital requirements.

3.1. Benchmark Analysis: Interbank Market Closed

The benchmark assumes no trading among banks ($\lambda = 0$ in (6)). Each bank's date 0 strategies consist of (i) public capital structure choice; (ii) private screening decision; and (iii) public loan approval/rejection decision. For a given θ, the equilibrium concept for the game between a bank and its financiers is subgame perfect equilibrium. The bank's capital structure informs financiers about its screening, based on which bank debt is priced. The bank's capital structure and screening choices maximize its expected net profit, anticipating the financiers' response to the chosen capital structure (which indicates whether it will screen as Holmstrom and Tirole 1997). It is a game of hidden action (will the bank screen?), not hidden type.

Because all banks are a priori identical, act independently, and follow the same decision rule, they will behave identically. That is, the equilibrium for a given θ will be *homogeneous* in the sense that all banks play the same strategies (i.e., make the same capital structure and screening choices). We show that the equilibrium for a given θ is unique and involves all banks either screening ("screening equilibrium") or avoiding screening ("no-screening equilibrium"). In the screening equilibrium, each bank holds enough equity to ensure that it has an incentive to screen, similar to the role of monitoring capital in Holmstrom and Tirole (1997). Each bank posts no equity in the no-screening equilibrium. There exists a cutoff value of θ, such that the equilibrium is a no-screening (respectively, screening) equilibrium for all θ values above (respectively, below) the cutoff. To get these results, we first examine the payoffs to banks when they all screen and when none screens, respectively, and then show how the specific equilibrium obtains for a given θ.

3.1.1. Screening. Suppose financiers believe the bank screens, so only the G loan will be approved at $t=0$. The date 2 repayment obligation on the legacy debt D, perceived to be risk free by financiers, is thus D. The following incentive-compatibility (IC) constraint must hold to ensure bank screening:

$$\begin{aligned} &-E-\psi E+0.5(X-D+X-I)+0.5E-c \\ &\geq \max\{-\psi E,\ -E-\psi E+[\theta+0.5(1-\theta)] \\ &\quad (X-D+X-I)-0.5(1-\theta)Y(\theta)\}. \end{aligned} \tag{7}$$

The left-hand side (LHS) of (7) is the bank's expected profit from screening, net of the equity input (E), and the cost of equity (ψE). The right-hand side (RHS) is the bank's expected net profit if it raises equity E but avoids screening, in which case the bank either approves or rejects the unscreened loan, depending on which choice yields higher profit. Rejection results in a net loss, ψE. Approving the unscreened loan will return X at $t=2$ if the macro state is a boom or a recession but the loan was G to begin with. Thus, the probability of the loan paying off X is $\theta+0.5(1-\theta)$. The bank's net profit in this case is $(X-D)+(X-I)$. The probability is $0.5(1-\theta)$ that an M loan was extended and there is a recession, in which case the loan returns zero and the bank defaults at $t=2$, losing its equity and charter value, $Y(\theta)$.

Both sides of (7) assume that financiers believe the bank with equity E has screened, so it solves for the minimum equity to induce bank screening when financiers believe it will screen: $E \geq E^{\text{hom}}(\theta)$,[18] where (the superscript "hom" indicates that all banks behave identically)

$$E^{\text{hom}}(\theta)=\frac{2\theta(X-I)+2c}{1-\theta}-Y(\theta). \tag{8}$$

Given costly equity, the bank holds $E=E^{\text{hom}}(\theta)$ in equilibrium to guarantee incentive compatibility.[19]

The bank's expected net profit with screening is (letting $E=E^{\text{hom}}(\theta)$ in the LHS of (7))

$$\pi_{\text{s}}^{\text{hom}}(\theta)=X-I-\psi E^{\text{hom}}(\theta)-c. \tag{9}$$

3.1.2. No Screening. Suppose financiers believe the bank avoids screening and lends unconditionally, so the legacy loan is equally likely to be G or M. Because M yields zero in a recession, the market's date 0 expectation is that the loan will return X with probability $\theta+0.5(1-\theta)$. The repayment obligation on the legacy debt is thus $\frac{D}{\theta+0.5(1-\theta)}=\frac{2D}{1+\theta}$. Moreover, equity is unnecessary absent screening, so the loan is entirely deposit funded ($D=I$), with a repayment obligation of $\frac{2I}{1+\theta}$.

If the date 1 macro state is a boom, the legacy loan will pay off X. Given $X>(2-\delta)I$ in (1), the bank can use the loan's date 2 cash flow X to repay its legacy debt and fund the new asset. If the macro state is a recession, the market believes the (unscreened) legacy loan will return X or zero with equal probability. The bank cannot fund the new asset in this case. The bank's pledgeable cash flow available for financing the new asset after repaying its (more senior) legacy debt is $0.5\left(X-\frac{2I}{1+\theta}\right)$. By (1), this is less than what is needed to fund the new asset, $(1-\delta)I$ (δI comes from the new asset's own pledgeable cash flow).

The bank's expected net profit without screening is

$$\begin{aligned} \pi_{\text{ns}}^{\text{hom}}(\theta)=&[\theta+0.5(1-\theta)]X-I+\theta(X-I) \\ &-0.5(1-\theta)Y(\theta). \end{aligned} \tag{10}$$

3.1.3. Equilibrium. For a given $\theta\in[0,\overline{\theta}]$, the previous analysis shows that a bank either raises equity $E^{\text{hom}}(\theta)$ and screens or posts no equity and avoids screening. If a bank screens, it raises exactly $E^{\text{hom}}(\theta)$, and its expected net profit is $\pi_{\text{s}}^{\text{hom}}(\theta)$ in (9). If the bank does not screen, it holds no equity,[20] and its expected net profit is $\pi_{\text{ns}}^{\text{hom}}(\theta)$ in (10). The (unique) equilibrium outcome for a given θ depends on how $\pi_{\text{s}}^{\text{hom}}(\theta)$ compares to $\pi_{\text{ns}}^{\text{hom}}(\theta)$. Because all banks are a priori identical and act independently absent trading, they play identical equilibrium strategies for any given θ.

Proposition 1 (Equilibrium with Interbank Market Closed). *Suppose the cost of screening c takes an intermediate* *value (see (A.1)). Then, there exists θ^* such that for each $\theta \in [0, \theta^*)$ the unique equilibrium involves all banks raising equity $E^{\text{hom}}(\theta) > 0$ and screening at $t = 0$, whereas for each $\theta \in [\theta^*, \overline{\theta}]$, the unique equilibrium involves no bank holding equity and screening.*

We refer to the equilibrium wherein all banks screen as the "benchmark screening equilibrium" and the one wherein none screens as the "benchmark no-screening equilibrium."[21] The intuition is that the marginal value of screening out M loans declines with θ because M pays off X in a boom anyway. Denote

$$\Delta\pi^{\text{hom}}(\theta) \equiv \pi_{\text{s}}^{\text{hom}}(\theta) - \pi_{\text{ns}}^{\text{hom}}(\theta) \tag{11}$$

as the value-added of screening relative to no screening. Proposition 1 shows $\Delta\pi^{\text{hom}}(\theta) > 0$ for $\theta \in [0, \theta^*)$, $\Delta\pi^{\text{hom}}(\theta^*) = 0$, whereas $\Delta\pi^{\text{hom}}(\theta) < 0$ for $\theta \in (\theta^*, \overline{\theta}]$. The proof of the proposition also shows $\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta} < 0$.

3.2. Main Analysis: Interbank Market Open

3.2.1. Motivation for Heterogeneity. The benchmark equilibria in which all banks behave identically is a natural consequence of having a priori identical banks with no trading, but the equilibria are inefficient, relative to both the first best and the second best with trading. In the first best, screening is observable, so banks keep no equity but still screen, and new assets can be funded in all states. In the benchmark screening equilibrium, M loans are screened out, enabling all banks to fund new assets in both a boom and a recession. The inefficiency is that *every* bank posts costly equity to guarantee screening incentives. In the benchmark no-screening equilibrium, every bank avoids costs of screening and equity, but the inefficiency is that *none* can finance the new asset in a recession. Thus, both benchmark equilibria are inefficient compared with the first best.

The benchmark equilibria are also inefficient relative to the second best when we allow trading (i.e., $\lambda > 0$ in (6)). With trading, there is an equilibrium for certain values of the boom probability θ in which some banks raise equity and screen, whereas others avoid equity and screening, that is, a priori identical banks pursue different equilibrium strategies. Such *endogenous* heterogeneity can bridge the extreme outcomes in the two homogeneous equilibria in the no-trading benchmark, thereby increasing welfare. For every θ for which a trading equilibrium exists, welfare is higher than in the benchmark equilibrium for that θ.

In a heterogeneous equilibrium, banks with high capital invest in screening at $t = 0$. This date 0 screening investment is reusable to some extent in screening other similar loans at $t = 1$. This reusability of information enables these high-capital banks to screen the legacy loans of no-capital banks and cherry pick those they find to be good. This allows high-capital banks to raise external financing to purchase the legacy loans of no-capital banks, even though no-capital banks themselves cannot directly raise market financing against the cash flows of their own legacy loans. Essentially, high-capital banks play a certification role and act as de facto intermediaries between no-capital banks and the market. High-capital banks are able to buy loans at a discount, which generates a trading profit to offset the costs of equity and screening.

Banks that choose to hold no capital and do not screen anticipate selling their loans to high-capital banks in a recession. Gains to trade arise from no-capital banks avoiding the costs of equity and screening yet financing new assets in a recession with a positive probability. The trading price splits these gains between the buyer and the seller. The equilibrium price is endogenously determined to ensure that the surplus sharing gives each high-capital bank and no-capital bank the same expected profit.

3.2.2. Equilibrium. The plan for this section is as follows. First, we determine the price, p, at which legacy loans will be traded, taking as given the masses of buyers and sellers (N_b and N_s, respectively), hence the market tightness $n = \frac{N_b}{N_s}$. Second, we show that trading strengthens screening incentives of buyers, so buyers can hold less equity while maintaining screening incentives. Third, we establish the existence and uniqueness of a heterogeneous equilibrium (for certain θ values) by solving for N_b and N_s and examine its properties. Finally, we compare the trading-induced heterogeneous equilibrium to the two no-trading benchmark equilibria that occur over different sets of θ values, and show that, for every θ, the heterogeneous equilibrium has higher social welfare than the (unique) benchmark equilibrium for that θ. We also show that relative to the social optimum, there are excessive buyers in the heterogeneous equilibrium when each bank freely chooses to be a buyer or seller.

We start with the presumption that some banks raise equity and screen at $t = 0$, whereas others do not, with the former being potential loan buyers and the latter being loan sellers at $t = 1$. We verify this later as an equilibrium outcome (see Endnote 22). In a boom, it is common knowledge that all banks' legacy loans will return X at $t = 2$, whether they were screened or not. This provides sufficient borrowing capacity to the bank at $t = 1$ to enable it to issue new debt to finance the new asset and repay its legacy debt at $t = 2$ as well. Consequently, there is no trading in a boom, and we focus our analysis on the date 1 macro state being a recession.

3.2.2.1. Loan Price. Suppose a buyer is matched with a seller and the buyer has verified the seller's loan is G. The buyer pays p to purchase the G loan, so its surplus from the trade is $X - p$. As shown earlier (see Endnote 14), there is no trade in a recession if the loan is identified to be M (which returns zero).

The seller also benefits from selling a G loan: not selling it leaves the seller unable to fund the new asset. This is because depositors who would need to fund the bank's new asset at $t = 1$ view it a fifty-fifty bet that the bank's legacy loan is G, and that does not provide a high enough pledgeable cash flow to support the necessary financing. Selling the G loan gives the seller the needed funds to invest in the new asset. The seller's surplus from trade, relative to no trade, is $(p - \tilde{D}) + (X - I) - (X - \tilde{D}) = p - I$, where $\tilde{D}$ denotes the seller's date 2 repayment obligation on its legacy debt.

Therefore, the total surplus from trade is $(X - p) + (p - I) = X - I$, which is exactly the surplus generated by the seller's new asset: surplus that would have been lost in the absence of trade. The buyer and the seller bargain over the price that will determine the sharing of this surplus. This price p is a solution to the following Nash bargaining problem:

$$\max_{p} \; (X - p)^{\beta}(p - I)^{1-\beta}, \tag{12}$$

where $\beta = \frac{1}{1+n}$ (see the parametrization in (6)) is the buyer's bargaining power. The solution is

$$p = X - \frac{X - I}{1 + n}. \tag{13}$$

Note $p \in [I, X]$ and is increasing in n, the market tightness (buyer/seller ratio). A higher n weakens the buyer's bargaining power and thus raises the price. The price discount is $\frac{X-I}{1+n}$, which compensates for the buyer's costs of screening and raising equity (to ensure screening). We can interpret p as the liquidation value of the seller's G legacy loan in a recession, which endogenously depends on n.

3.2.2.2. Trading and Screening. We show that the minimum equity needed to induce buyer screening for a given θ, denoted by $E^{\text{het}}(\theta)$ (the superscript "het" indicates that banks behave differently), is lower than that in the benchmark screening equilibrium, $E^{\text{hom}}(\theta)$ (see (8)). The intuition is that trading gains are available to a buyer only if the buyer has invested in (reusable) screening that enables it to cherry-pick a G legacy loan from a seller in a recession. The possibility of trading thus provides an additional screening incentive to the buyer, besides the screening-related benefit of avoiding making an M loan (by itself) at $t = 0$ that would imperil its charter value at $t = 2$. Less equity is needed therefore to incentivize screening.

Lemma 1 (Trading Induces Screening). *Trading increases a buyer's screening incentive and hence reduces the reliance on costly equity as skin in the game to induce buyer screening:*

$$E^{\text{het}}(\theta) = E^{\text{hom}}(\theta) - \lambda \frac{X - p}{1 + n}. \tag{14}$$

The reduction in the minimum buyer equity, $\lambda \frac{X-p}{1+n}$, is larger when the buyer/seller ratio n is smaller and/or the interbank market efficiency λ is higher.

Compared with $E^{\text{hom}}(\theta)$, $E^{\text{het}}(\theta)$ is smaller by $\lambda \frac{X-p}{1+n}$, which captures the extra screening incentive from a buyer hoping to correctly identify the quality of the seller's loan and purchase a G loan at a discounted price in a recession, instead of eschewing buying in the absence of screening.

In a recession, a buyer meets a seller with probability $\frac{\alpha(n)}{n} = \frac{\lambda}{1+n}$. The buyer's date 0 investment in screening allows it to identify the seller's loan type and only purchase a G loan, yielding a gain of $X - p$. If the buyer deviates and does not invest in screening at $t = 0$, then no trade occurs even if it is matched with a seller because the buyer lacks the ability to discern whether the seller's loan is G or M.[22] Therefore, the date 0 investment in screening enables the buyer to gain $\frac{\alpha(n)}{n}(X - p) = \lambda \frac{X-p}{1+n}$ from the date 1 trading in a recession, which is exactly the amount by which its equity is reduced.[23]

As the market tightness n declines or the interbank market efficiency λ improves, it becomes easier for a buyer to meet a seller (the probability $\frac{\alpha(n)}{n} = \frac{\lambda}{1+n}$ increases). A lower n further increases the buyer's bargaining power, so p decreases. Thus, the buyer's gain from trading, $\lambda \frac{X-p}{1+n}$, increases. This amplifies the buyer's potential loss from not investing in screening because, absent screening, the buyer forgoes more opportunities to buy a G loan at an even lower price as n declines or λ increases. The effect of trading on screening thus becomes stronger, leading to a bigger gap between $E^{\text{hom}}(\theta)$ and $E^{\text{het}}(\theta)$: for every θ, even less equity is needed to induce screening relative to that in the benchmark screening equilibrium.

It is useful to note that $E^{\text{het}}(\theta)$, derived from the IC constraint in (A.2) assuming the belief that a bank holding $E^{\text{het}}(\theta)$ has screened, is the *minimum* capital the bank needs to post to convince financiers of its screening investment. Suppose we have an equilibrium in which some banks hold $E^{\text{het}}(\theta)$ and screen (and become buyers), whereas others avoid equity and screening (and become sellers). Clearly, no bank would raise $E > E^{\text{het}}(\theta)$, which is a waste of equity. Suppose one bank deviates by choosing $E \in (0, E^{\text{het}}(\theta))$. Assume, counterfactually, that financiers believe this bank will still screen and be a buyer. It follows immediately that (A.2) will be violated and the bank will not screen (hence, cannot be a buyer), so such a belief is not rational. Moreover, it cannot be an equilibrium either that this bank becomes a seller, because posting equity (with or without screening) does not change the price at which the bank can sell its loan to a buyer (who screens) in a recession, so the equity and screening (if the seller were to screen) are just wasted. If a boom occurs, then no selling is needed, and the equity and screening are again wasted. In sum, it cannot be an equilibrium that any bank with equity less than $E^{\text{het}}(\theta)$ will screen. Therefore, in a heterogeneous equilibrium a bank either holds $E^{\text{het}}(\theta)$ and screens or completely avoids equity and screening.

As shown later, n and p, both of which are endogenously determined, depend on buyer equity, $E^{\text{het}}(\theta)$. There is thus an interesting feedback loop between equity and trading. The amount of equity needed to support screening affects a bank's decision to be a buyer or seller, so it impacts market tightness n, which then affects the loan price p. In turn, both n and p affect $E^{\text{het}}(\theta)$ (Lemma 1). The equilibrium characterization below shows how all these endogenous variables are driven by the boom probability (θ) and the interbank market efficiency (λ), which are two deep parameters of the model.

3.2.2.3. Equilibrium Characterization. Suppose θ is below θ^* characterized in Proposition 1. For all these θ values, when the interbank market is closed for trading, the benchmark equilibrium involves all banks raising equity and screening. When trading is allowed, suppose one bank deviates by holding no equity and avoiding screening, and plans to sell its loan at $t=1$. If the interbank market efficiency, λ, is very low, then the likelihood of trading is very low and the seller's expected gain from trade will be too low to cover its loss from the deviation, $\Delta\pi^{\text{hom}}(\theta)$ (see (11)). Therefore, the equilibrium involves all banks holding equity and screening, despite the availability of trading. Next, suppose λ is sufficiently high, so trading is quite likely. Because $\Delta\pi^{\text{hom}}(\theta)$ decreases with θ, a high enough θ makes it profitable for the bank to avoid screening and instead plan to sell its loan. Because the deviating bank is the only seller, the price at which it can sell its G legacy loan is close to the loan's true value X, so the deviating bank captures almost all the trade surplus. This induces more deviations, lowering p, and continues until the expected profits of buyers and (no-equity, no-screening) sellers are equal. This pins down the market tightness n, hence the loan price p (via (13)), and explains how a heterogeneous equilibrium arises when $\theta < \theta^*$ and λ is sufficiently high.

Formally, there is a unique cutoff $\theta_L \in (0, \theta^*)$, given by (A.4), such that when $\theta \in [0, \theta_L]$ the equilibrium has all banks behaving identically, raising equity and screening (as in the benchmark screening equilibrium), we call this a "homogeneous" equilibrium. When $\theta \in (\theta_L, \theta^*)$, some banks choose to avoid equity and screening, whereas others continue to hold equity and screen. The indifference condition establishing this heterogeneous equilibrium is in (A.5). Consistent with intuition, θ_L is lower when λ is higher: the trading-induced heterogeneous equilibrium is more likely to emerge with a more efficient interbank market.

Next, consider $\theta \in [\theta^*, \overline{\theta}]$. When trading is prohibited, the benchmark equilibrium involves all banks avoiding equity and screening. When trading is allowed, suppose one bank deviates by raising (enough) equity and screening at $t=0$, hoping to harvest gains from trade associated with buying a G loan at a discounted price in a recession. Because it is the only buyer, the loan price will be very low (close to I). The buyer extracts almost all the trade surplus. This induces more deviations, increasing p, and continues until the expected profits of buyers and sellers are equal. However, when the boom probability θ is very high or the interbank market efficiency λ is very low, the likelihood of trading is very low, causing the buyer's expected gain from trade insufficient to cover the costs of equity and screening. Therefore, with a high enough θ or a low enough λ, deviating from the benchmark no-screening equilibrium is not profitable.

Formally, there is a unique cutoff $\theta_H \in (\theta^*, \overline{\theta})$, given by (A.7), such that when $\theta \in [\theta_H, \overline{\theta}]$ the equilibrium has all banks behaving identically, avoiding equity and screening (as in the benchmark no-screening equilibrium), we again call this a homogeneous equilibrium. When $\theta \in [\theta^*, \theta_H)$, some banks choose to hold equity and screen, whereas others continue to avoid equity and screening. The indifference condition characterizing this heterogeneous equilibrium is in (A.8). Consistent with the intuition that a more efficient interbank market facilitates the trading-induced heterogeneous equilibrium, θ_H is higher when λ is higher.

We now state our main result, which characterizes the equilibrium when interbank trading is allowed, as well as its comparative statics properties with respect to the model's deep parameters, θ and λ.

Proposition 2 (Equilibrium with Interbank Market Open). *Suppose the cost of screening c takes an intermediate value (see (A.1)), and the value-added of screening is sufficiently big when $\theta = 0$ (see (A.3)) while sufficiently small when $\theta = \overline{\theta}$ (see (A.6)). There are cutoff values for θ, θ_L and θ_H, determined by (A.4) and (A.7), respectively. For each $\theta \in [0, \theta_L]$, the unique equilibrium is homogeneous with all banks raising equity $E^{\text{hom}}(\theta)$ and screening. For each $\theta \in [\theta_H, \overline{\theta}]$, the unique equilibrium is homogeneous with no bank posting equity and screening. For each $\theta \in (\theta_L, \theta_H)$, the unique equilibrium is heterogeneous with some banks (mass N_b) holding equity $E^{\text{het}}(\theta)$ and screening while the rest (mass N_s) avoiding equity and screening, and it has the following properties:*

1. *The interbank market tightness (buyer/seller ratio),* $n = \frac{N_b}{N_s}$*, is characterized by*

$$\lambda(X - I)\left[\frac{1-\theta}{2}\frac{n-1}{n+1} - \frac{\psi}{(n+1)^2}\right] = \Delta\pi^{\text{hom}}(\theta), \qquad (15)$$

with the equilibrium loan price $p = X - \frac{X-I}{1+n}$*. The region* (θ_L, θ_H) *expands on both ends as the interbank market efficiency* λ *increases.*

2. *Both* n *and* p *decrease with* θ *over the region* (θ_L, θ_H)*. When* λ *increases,* n *and* p *become less sensitive to* θ*, that is,* $|\frac{\partial n}{\partial \theta}|$ *and* $|\frac{\partial p}{\partial \theta}|$ *decrease with* λ*.*

For certain θ values, a heterogeneous equilibrium arises with trading, as heterogeneity improves the welfare of banks that keep equity and become loan buyers and the welfare of those that hold no equity and are sellers. The anticipation of gains to trade helps achieve the twin goals of economizing on equity and screening while also enabling investments in new assets in a recession. When a boom is more likely (higher θ), trading becomes less likely, so to compensate a buyer for its costs of screening and holding equity, the loan price p decreases, implying a lower buyer/seller ratio n. A more efficient interbank market (higher λ) dampens the effect of θ on trading opportunities, so p and n decline less when θ increases.

It is intuitive that the market tightness (n) falls as the boom probability (θ) increases and this result may be encountered in other settings as well. What is more interesting is that the *sensitivity* of n to θ declines as the interbank market efficiency λ increases. We show in Proposition 3 that the heterogeneous equilibrium has too many buyers (n too high) relative to the social optimum (which leads to excessive buyer equity). Therefore, the implication is that improving the efficiency of the interbank market for trading can reduce the inefficiency of the equilibrium involving excessive buyers (and their equity) during periods in which banks perceive a low boom probability.

3.2.2.4. Degree of Heterogeneity. The degree of heterogeneity among banks is highest when $n = 1$ (equal mass of buyers and sellers). When $n > 1$, there is less heterogeneity when n increases (more and more buyers), whereas when $n < 1$, the degree of heterogeneity decreases when n decreases (more and more sellers). Because $\frac{\partial n}{\partial \theta} < 0$ (Proposition 2), as θ increases over the region (θ_L, θ_H), the degree of heterogeneity first increases and then decreases, peaking at some $\theta = \theta^{\max}$.[24]

3.3. Welfare and Equilibrium Efficiency

In this section, we compare the competitive equilibrium derived in Proposition 2 to what a social planner would implement. First, we show that when trading is permitted, switches of the economy from a homogeneous screening equilibrium to a heterogeneous equilibrium at $\theta = \theta_L$ and from a homogeneous no-screening equilibrium to a heterogeneous equilibrium at $\theta = \theta_H$ are socially optimal. Second, compared with the social planner's optimum, the heterogeneous equilibrium for each $\theta \in (\theta_L, \theta_H)$ is nonetheless inefficient.

To understand the first result, note that the cutoffs θ_L and θ_H are determined in the same way (see (A.4) and (A.7)) that a social planner would. When $\theta \le \theta_L$, no individual bank wants to deviate from the benchmark screening equilibrium and become the only seller because both the social surplus from trading and the associated social loss from avoiding screening are *fully* internalized by the seller, and the surplus fails to make up for the loss. Similarly, no individual bank deviates from the benchmark no-screening equilibrium to be the only buyer when $\theta \ge \theta_H$ because the buyer *fully* internalizes both the trading surplus and the associated screening cost, and the surplus does not compensate for the cost.

To see the equilibrium inefficiency for $\theta \in (\theta_L, \theta_H)$, it is useful to first preview the idea. Lemma 1 shows that trading provides a buyer an extra incentive to screen, which is socially beneficial because it reduces the reliance on equity. This incentive effect of trading is stronger with a lower market tightness n, because a lower n increases the probability that a buyer meets a seller, and also lowers the loan price conditional on trading. Both contribute to increase the buyer's potential gain from trading, making screening more valuable to the buyer, thereby further reducing the equity needed to incentivize screening.

However, no individual bank, when making its own decision to be a buyer or seller, *fully* internalizes this social value of a lower n. Said differently, each atomistic buyer ignores the negative externality it imposes on other buyers because an elevated n results in *all* buyers posting more equity. As a result, the privately optimal heterogeneous equilibrium characterized in Proposition 2 has excessive market tightness and bank equity, relative to what a social planner would prefer.

It is worthwhile pointing out that this externality differs from the standard congestion externality in the literature. Under a congestion externality, greater buyer entry in the interbank market lowers the matching probability of each buyer with a seller. This externality is ruled out by the Hosios condition (we specify a buyer's bargaining power as $\beta = \frac{n\alpha'(n)}{\alpha(n)}$), under which banks fully internalize search externalities, resulting in efficient buyer-seller matching. Turning off the congestion externality channel allows us to focus on the new externality described previously; although buyers do not generate any congestion in matching, they fail to internalize the negative equity externality imposed on others.

To formalize this, we conduct a welfare analysis. Banks extract all surplus, so computing social welfare is equivalent to computing bank surplus. The social welfare in a benchmark equilibrium for a given θ is

$$W^{\mathrm{hom}}(\theta) = \begin{cases} N\pi_{\mathrm{s}}^{\mathrm{hom}}(\theta) & \text{for } \theta \in [0, \theta^{*}) \\ N\pi_{\mathrm{ns}}^{\mathrm{hom}}(\theta) & \text{for } \theta \in [\theta^{*}, \overline{\theta}], \end{cases} \quad (16)$$

where $\pi_{\mathrm{s}}^{\mathrm{hom}}(\theta)$ and $\pi_{\mathrm{ns}}^{\mathrm{hom}}(\theta)$ are given by (9) and (10), respectively. When trading is permitted, although the equilibrium remains homogeneous for $\theta \in [0, \theta_L] \cup [\theta_H, \overline{\theta}]$ with welfare $W^{\mathrm{hom}}(\theta)$, a heterogeneous equilibrium arises for $\theta \in (\theta_L, \theta_H)$, generating welfare $W^{\mathrm{het}}(\theta, \lambda, n) = W^{\mathrm{hom}}(\theta) + \Delta W(\theta, \lambda, n)$, where

$$\Delta W(\theta, \lambda, n) = \begin{cases} N_s[0.5(1-\theta)\alpha(n)(X-I)] \\ \quad + N_b\psi\dfrac{\alpha(n)}{n}(X-p) - N_s\Delta\pi^{\mathrm{hom}}(\theta) & \text{for } \theta \in (\theta_L, \theta^{*}) \\ N_s[0.5(1-\theta)\alpha(n)(X-I)] + N_b\psi\dfrac{\alpha(n)}{n}(X-p) \\ \quad + N_b\Delta\pi^{\mathrm{hom}}(\theta) & \text{for } \theta \in [\theta^{*}, \theta_H). \end{cases} \quad (17)$$

Consider $\theta \in (\theta_L, \theta^{*})$. In the first term, the expression in the bracket is a seller's new asset surplus captured because of trade (the price p is a pure transfer with no welfare effect); N_s is the mass of sellers. The second term, $\psi\frac{\alpha(n)}{n}(X-p)$, is each buyer's equity cost reduction because of the incentive effect of trading; N_b is the mass of buyers. As the preceding discussion shows, this is the key externality explored. The third term, $N_s\Delta\pi^{\mathrm{hom}}(\theta)$, is the sellers' aggregate profit loss because of switching from screening (in the benchmark screening equilibrium) to no screening (in the heterogeneous equilibrium). For $\theta \in [\theta^{*}, \theta_H)$, $\Delta W(\theta, \lambda, n)$ only differs in the last term, $N_b\Delta\pi^{\mathrm{hom}}(\theta)$, aggregate profits lost by buyers who deviate from no-screening (in the benchmark no-screening equilibrium) to screening (in the heterogeneous equilibrium).[25]

We show $\Delta W(\theta, \lambda, n) > 0 \;\; \forall\theta \in (\theta_L, \theta_H)$, and the trading-induced welfare enhancement $\Delta W(\theta, \lambda, n)$ increases with market efficiency λ.[26] Nonetheless, for each $\theta \in (\theta_L, \theta_H)$, we prove that the privately optimal n determined by (15) exceeds the social planner's n, which is the one maximizing $\Delta W(\theta, \lambda, n)$ for that θ.

Proposition 3 (Welfare and Equilibrium Efficiency). *When trading is allowed, the equilibrium switches at $\theta \in \{\theta_L, \theta_H\}$ are socially optimal. Nonetheless, the heterogeneous equilibrium for every $\theta \in (\theta_L, \theta_H)$ is inefficient relative to what a social planner would implement; compared with the social optimum, the equilibrium involves too many buyers relative to sellers, too high a loan price, and excessive buyer equity.*

It is interesting that the welfare distortion here is the opposite of that in fire sales models. With fire sales, the asset price is too low (selling assets to nonexpert second-best users), which causes bank failures because sharp declines in asset prices wipe out bank equity. Thus, even otherwise healthy banks suffer because of fire sales of common assets by other banks. Here, the distortion is that the price is too high relative to the social optimum, and the underlying economic mechanism is completely different.

The price being "too high" is not in itself an inefficiency per se in our model, as the price is a pure transfer. Unlike fire sales models in which low prices cause inefficient transfers, here sellers and buyers are equally skilled in managing assets, so the sale does not generate an inefficiency. Rather, the inefficiency is that the high price reflects an excessive number of buyers, with none internalizing the general equilibrium inefficiency of buyers in the aggregate keeping too much equity relative to the first best.

Although this distortion stands in contrast to the usual distortion of excessive leverage in banking models, it should be remembered that our result obtains in a setting devoid of deposit insurance or other safety nets. As we know from Merton (1977) and the subsequent literature, these safety nets induce banks to be excessively leveraged. What we show is that in unregulated banking (no safety nets) with trading in opaque over-the-counter markets, excessive information production, a sort of "arms race" in information acquisition among buyers, is a possible distortion. This is somewhat related to the point made by Bolton et al. (2016a), namely that the financial sector produces too much information. In their model, this happens in the context of excessive rent extraction by the financial sector. The forces at play in our model are different: The excessive information production is associated with buyers keeping aggregate capital that is excessive relative to the social optimum. However, the (heterogeneous) equilibrium also involves sellers keeping zero capital, so the excessive capital result applies only to buyers.

4. Robustness, Extension, and Implications

4.1. Robustness and Extension

We have assumed $\kappa = 0$ (information is fully reusable) in the main analysis. In Appendix B.1, we analyze the model with $\kappa > 0$ (information is only partially reusable) and prove the following result.

Proposition 4 (Information Reusability). *The results in the main analysis hold qualitatively as long as κ is not too large, but there will be no heterogeneous equilibrium if κ is too high.*

This result highlights the key role of information reusability. An interesting, and perhaps somewhat surprising, implication is that heterogeneity is more likely to emerge among banks intermediating in industries with greater similarity among firms and higher cross-sectional information reusability (lower κ).

In our main analysis, banks choose privately optimal capital ratios, and there is no social externality that necessitates a regulatory capital requirement.[27] Nonetheless, because of the prominent role played by bank capital, it is interesting to ask what the effect of such a requirement would be. Assuming a minimum capital requirement $E^{\min}$, we provide an analysis on this in Appendix B.2 with the following finding.

Proposition 5 (Capital Requirements). *The heterogeneous equilibrium will not emerge if $E^{\min}$ is sufficiently high. This effect of $E^{\min}$ becomes stronger at the margin with a higher λ.*

Appendix B.2 shows that a higher $E^{\min}$ shrinks the heterogeneous equilibrium region (θ_L, θ_H) and shrinks it faster when market efficiency λ is higher. The takeaway is that regulators need to balance this adverse effect of capital requirements on the (welfare-enhancing) heterogeneous equilibrium against other social benefits (e.g., reduction of systemic risk, and enhanced depositor welfare) that motivate capital requirements in the first place. This tension becomes more pertinent with a more efficient interbank trading market. The intuition is that when the interbank market is more efficient, capital requirements cause banks to lose more at the margin from the loss of trading caused by the disappearance of the heterogeneous equilibrium. This suggests that, holding the social benefit of capital fixed, the socially optimal capital requirement would be lower with a more efficient interbank market.

4.2. Policy Implications

4.2.1. Prudential Regulation and Bank Capital Requirements. Our analysis highlights that even a banking system that starts out being homogeneous may transform itself into one that is heterogeneous to exploit potential gains from trade (Proposition 2). Comparing the benchmark (no trading) and heterogeneous (trading) equilibria, two interesting observations emerge. First, fewer banks hold equity in the heterogeneous equilibrium compared with the benchmark screening equilibrium in which all banks hold equity. Second, for every θ (boom probability), even those banks that hold equity in the heterogeneous equilibrium hold less equity than the banks holding equity in the benchmark screening equilibrium (Lemma 1). Therefore, the banking system as a whole raises less capital in the heterogeneous equilibrium than in the benchmark screening equilibrium and yet obtains higher welfare. This difference becomes larger with higher interbank market efficiency.

A lot of prudential bank regulation has the effect of establishing uniform standards of conduct and homogenizing banks. In particular, capital requirements, especially the Basel Accords, seek to establish a level playing field across banks in different countries, thus making banks more similar on this critical dimension.[28] Our analysis yields the somewhat surprising result that capital requirements can cause an increase in capital even in banks for whom the requirements are not binding. Sufficiently high capital requirements may eliminate the heterogeneous equilibrium and reduce welfare, and this issue is particularly concerning with a more efficient interbank market for asset trading (Proposition 5). This raises a novel issue for regulators to consider. In most discussions of capital requirements, the tension is between the social benefits of greater banking stability because of higher capital requirements and the private costs for banks in maintaining higher capital ratios (Thakor 2014). However, our analysis shows that what matters is not just the level of capital in banking but also the *distribution* of capital across banks, and regulators must weigh the financial stability benefits of higher capital requirements against welfare loss because of the elimination of the heterogeneous equilibrium.

Our analysis also shows that trading can induce bank screening (Lemma 1). Therefore, reducing interbank market search frictions (improving λ) may be a cost-efficient way to incentivize screening. This suggests a link between asset markets and bank regulation.

4.2.2. Fire Sales. Our analysis also highlights how bank heterogeneity may help avoid fire sales. Unlike the classic fire-sales model in which all firms are in the same boat, so assets end up with second-best users, bank heterogeneity facilitates trading among first-best users. The policy implication of this for central bank is significant, because preventing fire sales in financial markets is often a rationale for government bailouts. Central bank intervention to bail out failing institutions is sometimes justified on the ground that allowing an institution to fail may generate fire sales of common assets and hurt other institutions.[29] Thus, the government may intervene and bail out even nondepository institutions like Bear Stearns where bank runs are not an issue. Our analysis suggests that encouraging bank heterogeneity and reducing frictions in the secondary market for loan trading may reduce the need for bailouts.

4.2.3. Interbank Trading Markets. One of the inefficiencies in the heterogeneous equilibrium is that there are too many buyers relative to the social optimum when an interbank market for loan trading exists (Proposition 3). One way that policymakers can attempt to reduce this inefficiency is to tax the trading profits of buying banks (i.e., impose a Tobin's tax on financial transactions). Subsidizing selling banks would be another option, but that is likely to be inferior because it involves using taxpayer money.[30]

4.3. Empirical Implications

Our analysis has numerous empirical implications, many of which have been discussed earlier, but we consolidate here. First, consistent with Pérignon et al. (2018) and others, we show that when banks face funding stresses because of heightened insolvency concerns, as during the 2007–2009 crisis, not all banks will be frozen out of the credit market. Rather, high-capital banks will increase their short-term (uninsured) funding access and grow assets, whereas low-capital banks will retrench. In other words, crises will induce a reallocation of liquidity rather than a market-wide freeze.

Second, this reallocation will be achieved through active asset trading among banks, in which high-capital banks purchase assets from low-capital banks, consistent with the evidence in Irani and Meisenzahl (2017) that banks manage liquidity via loan sales during periods of stress and the finding in Berger and Bouwman (2013) that high-capital banks grow in size and market share during crises by purchasing assets from low-capital banks (see also Bord et al. 2021).

Third, our analysis predicts that even after controlling for bank size, asset composition, and so on, there will be capital structure heterogeneity among banks when they anticipate trading among themselves in the interbank loan market. There is ample empirical evidence that banks have different capital structures, even after controlling for observable differences like size, and that this generates different alphas (Bouwman et al. 2018). There are many possible reasons. One is the heterogeneous impact of regulatory supervision on bank risk and performance (Eisenbach et al. 2019, Kovner and Van Tassel 2019). Hirtle et al. (2019) show that even banks of the same size and with similar risk profiles receive different levels of supervision simply based on the Federal Reserve district they are headquartered in. When combined with the finding of Kovner and Van Tassel (2019) that supervision affects banks' cost of capital, heterogeneous supervisory practices by themselves can be a source of heterogeneity in banks' cost of capital, although not the only one. We show that the observed capital structure heterogeneity can arise even without any other (observable) differences in bank attributes or their costs of capital. Nonetheless, this heterogeneity then leads to differences in costs of capital as well as access to short-term funding during times of stress.

Fourth, given the different factors that can lead to bank heterogeneity noted previously, our analysis offers a prediction that can be used to distinguish our theory from other explanations for heterogeneity. Specifically, we predict that, controlling for all the other factors, higher trading efficiency in the interbank loan market generates more heterogeneity. International data may be useful in such a test.

Fifth, if all bank charter values increase (i.e., $Y(\theta)$ shifts up for all θ values), then banks will need to keep lower capital ratios to ensure screening incentives; see (8) and (14). Given capital requirements in practice, the prediction is that there will be less excess capital (i.e., capital in excess of regulatory requirements) with higher bank charter values. Therefore, in assessing bank fragility, we should consider not only capital ratios but also charter values: low capital ratios do not necessarily imply fragility if bank charter values are high.

5. Conclusion

We have developed a theory in which a priori identical banks may become heterogeneous in their capital structure and screening choices. We show that an equilibrium with such heterogeneity arises endogenously because it leads to trading that helps banks reduce inefficiencies associated with either excessive equity and screening or insufficient new asset investments by banks in the benchmark (no trading) equilibrium.

With heterogeneity, some banks, as the first-best users of banking assets, build financial muscle with sufficiently high precrisis capital and screening capability to put themselves in a position to purchase loans from banks with low capital. This can prevent fire sales and liquidity dry-ups in economic downturns. When a crisis arrives, high-capital banks have better legacy assets and continue to have access to funding, but low-capital banks with poorer assets are frozen out of the credit market. High-capital banks purchase legacy assets from low-capital banks and fund these purchases by increasing their short-term funding, whereas low-capital banks reduce their short-term funding. Thus, consistent with the recent empirical evidence, there is a reallocation of market liquidity from low-capital to high-capital banks without asset fire sales. It is also consistent with the empirical evidence that during financial crises, high-capital banks gain market share from low-capital banks (Berger and Bouwman 2013). According to this empirical evidence, the capital structure heterogeneity that enhances welfare in our model is observed with deposit-insured institutions. More recent evidence indicates that this heterogeneity is even greater among shadow banks (Jiang et al. 2020).

The banking literature has pointed out that relationship banking and bank specialization are both driven by information reusability. Our analysis echoes this by highlighting a central role played by information reusability in generating *future* expected trading surplus and, thereby, giving rise to ex ante heterogeneity among (a priori identical) banks in their investments in equity and screening.

Acknowledgments
The helpful comments of Allen Berger, Jason Donaldson, Ben Lester, Yao Zeng (discussant), and other participants in the Financial Intermediation Research Society 2021 conference and seminars at Washington University in St. Louis, Office of Financial Research, Federal Reserve Bank of Philadelphia, University of Nebraska, Villanova University, Chinese University of Hong Kong, and particularly those of two anonymous referees, an associate editor, and the editor (Bruno Biais) are gratefully acknowledged.

Appendix A. Proofs of Results in Section 3

Proof of Proposition 1. We first show (in details) how $E \geq E^{\text{hom}}(\theta)$ solves (7). Using $D = I - E$, the LHS of (7) can be simplified as $X - I - \psi E - c$, which strictly exceeds $-\psi E$ (the first term on the RHS), because $X - I - c > 0$ (following from (3) and (4)). Therefore, what is relevant for solving (7) is to compare its LHS with the second term on its RHS, which can be simplified as $-\psi E - 0.5(1-\theta)[E + Y(\theta)] + (1+\theta)(X - I)$ (using $D = I - E$). This leads to $E \geq E^{\text{hom}}(\theta)$, given by (8). We also verify $-\psi E - 0.5(1-\theta)[E + Y(\theta)] + (1+\theta)(X-I) > -\psi E$ in the neighborhood of $E = E^{\text{hom}}(\theta)$, so the first term on the RHS of (7), $-\psi E$, is indeed irrelevant for solving the IC constraint (see discussions in Endnote 18). Evaluating the LHS of the previous inequality at $E = E^{\text{hom}}(\theta)$ results in $X - I - \psi E^{\text{hom}}(\theta) - c$. To show this to be positive, given $\frac{dE^{\text{hom}}(\theta)}{d\theta} < 0$, it is sufficient to show $X - I - \psi E^{\text{hom}}(0) - c = X - I - c - \psi[2c - Y(0)] > 0$, which is exactly the condition in (4). The partial derivative of the LHS with respect to E is $-\psi - 0.5(1-\theta)$, whereas the partial derivative of the RHS with respect to E is $-\psi$; the LHS thus increases faster when E falls. Therefore, there exists $E' > E^{\text{hom}}(\theta)$, such that (by continuity) $-\psi E - 0.5(1-\theta)[E + Y(\theta)] + (1+\theta)(X-I) > 0 > -\psi E \ \forall E \in [0, E')$.

We now prove the rest of the proposition. Let $\Delta\pi^{\text{hom}}(\theta) \equiv \pi_{\text{s}}^{\text{hom}}(\theta) - \pi_{\text{ns}}^{\text{hom}}(\theta) = 0.5(1-\theta)X - \theta(X-I) + 0.5(1-\theta)Y(\theta) - \psi E^{\text{hom}}(\theta) - c$. We have $\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta} = -1.5X - 0.5Y(\theta) + \psi\left[Y'(\theta) - \frac{2(X-I+c)}{(1-\theta)^2}\right] + 0.5(1-\theta)Y'(\theta)$, which is strictly negative if ψ is not too big; a sufficient condition for this is given by (5). Thus, the existence and uniqueness of the cutoff $\theta^* \in (0, \overline{\theta})$, such that $\Delta\pi^{\text{hom}}(\theta) > 0$ for $\theta \in [0, \theta^*)$, $\Delta\pi^{\text{hom}}(\theta^*) = 0$, and $\Delta\pi^{\text{hom}}(\theta) < 0$ for $\theta \in (\theta^*, \overline{\theta}]$, can be ensured by $\Delta\pi^{\text{hom}}(0) > 0$ and $\Delta\pi^{\text{hom}}(\overline{\theta}) < 0$, which can be expressed as

$$\frac{0.5(1-\overline{\theta})X}{1 + 2\psi(1-\overline{\theta})^{-1}} - \overline{\theta}(X-I) + 0.5(1-\overline{\theta})Y(\overline{\theta}) < c < \frac{0.5X}{1+2\psi} + 0.5Y(0). \tag{A.1}$$

For $\theta \in [0, \theta^*)$, given $\Delta\pi^{\text{hom}}(\theta) > 0$, the bank holds equity $E^{\text{hom}}(\theta)$ and screens, earning $\pi_{\text{s}}^{\text{hom}}(\theta)$. For $\theta \in (\theta^*, \overline{\theta}]$, given $\Delta\pi^{\text{hom}}(\theta) < 0$, the bank holds no equity and does not screen, earning $\pi_{\text{ns}}^{\text{hom}}(\theta)$. As shown in the text preceding the proposition, no bank would deviate, and the equilibrium for a given θ is unique. Finally, it is clear that all banks, a priori identical and without interaction, play the same equilibrium strategies for each θ. □

Proof of Lemma 1. With trading, the buyer's IC constraint for screening (for given n and p) is

$$\begin{aligned} &-E - \psi E + 0.5(X - D + X - I) + 0.5E + 0.5(1-\theta)\frac{\alpha(n)}{n}(X-p) - c \\ &\geq \max\{-\psi E,\ -E - \psi E + [\theta + 0.5(1-\theta)](X - D + X - I) \\ &- 0.5(1-\theta)Y(\theta)\}. \end{aligned} \tag{A.2}$$

Simplifying it and using $\alpha(n) = \frac{\lambda n}{1+n}$ lead to (14).[31] Both sides of (A.2) are computed based on the financiers' belief that a buyer holding equity E has invested in screening.[32] Therefore, a bank with equity less than $E^{\text{het}}(\theta)$ would not screen and become a buyer, even if it were believed to have invested in screening; this means that $E^{\text{het}}(\theta)$ is the minimum equity required to sustain screening for a buyer. Moreover, it cannot be an equilibrium either for a bank with equity $E \in (0, E^{\text{het}}(\theta))$ to be a seller: regardless whether the seller screens with $E > 0$, the loan price in a recession is still given by p in (13) because the buyer screens and can identify the seller's loan type, so the equity and screening (if the seller were to screen) are wasted. There will be no trading in a boom, in which case the seller's equity is again wasted. Therefore, the equilibrium we characterize, in which buyers hold exactly $E^{\text{het}}(\theta)$ and screen while sellers completely avoid equity and screening, is unique. Finally, it is clear that the buyer's equity reduction, $\lambda\frac{X-p}{1+n}$, is decreasing in n (note p is increasing in n; see (13)) but increasing in λ. □

Proof of Proposition 2. The proof assumes that the condition in (A.1) holds, so the cutoff $\theta^* \in (0, \overline{\theta})$ exists and is unique (as in Proposition 1). First, we characterize θ_L. For a seller, switching from screening (and holding positive equity) to no-screening (and holding no equity) causes its profit to change by $-\Delta\pi^{\text{hom}}(\theta)$ without trading (note $-\Delta\pi^{\text{hom}}(\theta) < 0$ for $\theta < \theta^*$; Proposition 1). In a recession (with probability $1-\theta$), the deviating bank, being the only seller ($n \uparrow \infty$), meets a buyer with probability λ (the probability $\alpha(n) \uparrow \lambda$ as $n \uparrow \infty$). After meeting, the (monopolistic) seller sells its loan with probability 0.5 (as its unscreened loan is G with probability 0.5) with a price $p = X$ (see (13) with $n \uparrow \infty$). Trading yields the seller a gain of $X - I$. Therefore, the deviating bank's net profit change is $-\Delta\pi^{\text{hom}}(\theta) + 0.5\lambda(1-\theta)(X-I)$, which is monotonically increasing in θ if $\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta} < -0.5\lambda(X-I)$; a sufficient condition for this is given by (5). This expression is positive when $\theta = \theta^*$, as $\Delta\pi^{\text{hom}}(\theta^*) = 0$. Suppose exogenous parameters are such that this expression is negative when $\theta = 0$:

$$\Delta\pi^{\text{hom}}(0) > 0.5\lambda(X-I), \tag{A.3}$$

then $\theta_L \in (0, \theta^*)$ is uniquely pinned down by

$$-\Delta\pi^{\text{hom}}(\theta_L) + 0.5\lambda(1-\theta_L)(X-I) = 0. \tag{A.4}$$

It is clear that (i) $\theta_L < \theta^*$, because $\Delta\pi^{\text{hom}}(\theta_L) > 0$ (from (A.4)), $\Delta\pi^{\text{hom}}(\theta^*) = 0$, and $\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta} < 0$; and (ii) θ_L is decreasing in λ, because the LHS of (A.4) is increasing in both θ_L and λ.

When $\theta \in (\theta_L, \theta^*)$, the indifference condition that establishes the heterogeneous equilibrium is

$$\underbrace{-\Delta\pi^{\text{hom}}(\theta) + 0.5(1-\theta)\alpha(n)(p-I)}_{\text{seller's gain}} = \underbrace{0.5(1-\theta)\frac{\alpha(n)}{n}(X-p) + \psi\frac{\alpha(n)}{n}(X-p)}_{\text{buyer's gain}}. \tag{A.5}$$

The LHS captures a seller's profit change from deviating from the benchmark screening equilibrium when trading is possible. Without trading, deviation results in a profit drop by $\Delta\pi^{\text{hom}}(\theta)$ (note $\Delta\pi^{\text{hom}}(\theta) > 0$ for $\theta < \theta^*$; Proposition 1). The second term is the seller's gain from trade to compensate for that loss. A seller meets a buyer in a recession (with joint probability $(1-\theta)\alpha(n)$), sells its loan with probability 0.5 (as the seller's unscreened loan is G with probability 0.5), and gains $p - I$ (see discussions preceding (13) for how the seller's surplus $p - I$ is determined). The RHS computes a buyer's profit increase relative to the benchmark screening equilibrium. A buyer meets a seller in a recession (with joint probability $(1-\theta)\frac{\alpha(n)}{n}$), purchases the seller's loan with probability 0.5 (as the seller's loan is G with probability 0.5), and gains $X - p$. Furthermore, as Lemma 1 shows, with trading the buyer's equity reduces by $\frac{\alpha(n)}{n}(X-p)$, so the second term on the RHS captures the reduced equity cost to the buyer.

Next, we characterize θ_H. For a buyer, switching from no-screening (and holding no equity) to screening (and holding positive equity) causes its profit to change by $\Delta\pi^{\text{hom}}(\theta)$ without trading (note $\Delta\pi^{\text{hom}}(\theta) < 0$ for $\theta > \theta^*$; Proposition 1). In a recession (with probability $1-\theta$), the deviating bank, being the only buyer ($n \downarrow 0$), meets a seller with probability λ (the probability $\frac{\alpha(n)}{n} \uparrow \lambda$ as $n \downarrow 0$). After meeting, the (monopolistic) buyer purchases the seller's loan with probability 0.5 (as the seller's unscreened loan is G with probability 0.5) with a price $p = I$ (see (13) with $n \downarrow 0$). Trading yields the buyer a gain of $X - I$. Trading also lowers the buyer's equity cost by $\psi\frac{\alpha(n)}{n}(X-p)$ (Lemma 1), which is $\psi\lambda(X-I)$ with $n \downarrow 0$ and $p \downarrow I$. Therefore, the deviating bank's net profit change is $\Delta\pi^{\text{hom}}(\theta) + 0.5\lambda(1-\theta)(X-I) + \psi\lambda(X-I)$, which is monotonically decreasing in θ, as $\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta} < 0$ (see the proof of Proposition 1). This expression is positive when $\theta = \theta^*$, as $\Delta\pi^{\text{hom}}(\theta^*) = 0$. Suppose exogenous parameters are such that this expression is negative when $\theta = \overline{\theta}$:

$$\Delta\pi^{\text{hom}}(\overline{\theta}) < -0.5\lambda(1-\overline{\theta})(X-I) + \psi\lambda(X-I), \tag{A.6}$$

then $\theta_H \in (\theta^*, \overline{\theta})$ is uniquely pinned down by

$$\Delta\pi^{\text{hom}}(\theta_H) + 0.5\lambda(1-\theta_H)(X-I) + \psi\lambda(X-I) = 0. \tag{A.7}$$

It is clear that (i) $\theta_H > \theta^*$, because $\Delta\pi^{\text{hom}}(\theta_H) < 0$ (from (A.7)), $\Delta\pi^{\text{hom}}(\theta^*) = 0$, and $\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta} < 0$; and (ii) θ_H is increasing in λ, because the LHS of (A.7) is decreasing in θ_H while increasing in λ.

The indifference condition that characterizes the heterogeneous equilibrium for $\theta \in [\theta^*, \theta_H)$ is

$$\underbrace{0.5(1-\theta)\alpha(n)(p-I)}_{\text{seller's gain}} = \underbrace{\Delta\pi^{\text{hom}}(\theta) + 0.5(1-\theta)\frac{\alpha(n)}{n}(X-p) + \psi\frac{\alpha(n)}{n}(X-p)}_{\text{buyer's gain}}. \tag{A.8}$$

The RHS and LHS capture the buyer's and the seller's gains in the heterogeneous equilibrium relative to the benchmark no-screening equilibrium, respectively. This condition has a similar explanation as that for (A.5), except for $\theta \in [\theta^*, \theta_H)$, it is the buyer deviating from the benchmark equilibrium, so the associated profit drop, $\Delta\pi^{\text{hom}}(\theta)$, now appears on the RHS (note $\Delta\pi^{\text{hom}}(\theta) < 0$ for $\theta > \theta^*$; Proposition 1).

It is clear from the previous proof that $0 < \theta_L < \theta_H < \overline{\theta}$, where the first and last inequalities are ensured by (A.3) and (A.6), respectively, and $\theta_L < \theta_H$ follows from the (proved) facts that $\theta_L < \theta^*$ while $\theta_H > \theta^*$.

The previous proof shows that banks will deviate from the benchmark equilibrium and play different strategies for $\theta \in (\theta_L, \theta_H)$. The consequent interbank trading strictly improves each bank's profit, leading to a heterogeneous equilibrium. Algebra reveals that the indifference conditions in (A.5) and (A.8) share the same mathematical form

$$\alpha(n)\left[\frac{1-\theta}{2}\left(p - I - \frac{X-p}{n}\right) - \psi\frac{X-p}{n}\right] = \Delta\pi^{\text{hom}}(\theta), \tag{A.9}$$

which is a unified characterization of the heterogeneous equilibrium for each $\theta \in (\theta_L, \theta^*) \cup [\theta^*, \theta_H)$. Substituting $\alpha(n) = \frac{\lambda n}{1+n}$ and $p = X - \frac{X-I}{1+n}$ into the previous condition leads to the equilibrium condition in (15). The LHS of (15) is monotonically increasing in n, whereas the RHS, $\Delta\pi^{\text{hom}}(\theta)$, is not a function of n. Therefore, the heterogeneous equilibrium for a given $\theta \in (\theta_L, \theta_H)$, characterized by n and p, exists and is unique.

Finally, we prove the comparative statics results. Apply the implicit function theorem to (15):

$$\frac{\partial n}{\partial \theta} = \frac{\frac{1}{2}\frac{n-1}{n+1} + \frac{1}{\lambda(X-I)}\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta}}{\frac{1-\theta}{(n+1)^2} + \frac{2\psi}{(n+1)^3}}. \tag{A.10}$$

To have $\frac{\partial n}{\partial \theta} < 0$, the numerator needs to be negative, a sufficient condition for which is $\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta} < -0.5\lambda(X-I)$, which is ensured by the condition in (5). Thus, $|\frac{\partial n}{\partial \theta}| = \frac{-\frac{1}{2}\frac{n-1}{n+1} - \frac{1}{\lambda(X-I)}\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta}}{\frac{1-\theta}{(n+1)^2} + \frac{2\psi}{(n+1)^3}}$; it is clear that when λ increases, $-\frac{1}{\lambda(X-I)}\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta}$ decreases, so $|\frac{\partial n}{\partial \theta}|$ falls. Because p is increasing in n, it follows that $\frac{\partial p}{\partial \theta} < 0$ and $|\frac{\partial p}{\partial \theta}|$ is decreasing in λ as well. □

Proof of Proposition 3. We start by proving the first part of the proposition. When $\theta = \theta_L$, we have $N_b = N$, $N_s = 0$, $n = \frac{N_b}{N_s} = \infty$, so $\frac{\alpha(n)}{n} = 0$. Therefore, $\Delta W(\theta_L, \lambda, n) = 0$. Together with the arguments preceding (A.4) regarding why no single bank, fully internalizing all social benefits and costs from trading, finds it profitable to deviate from the benchmark screening equilibrium when $\theta \le \theta_L$, this proves that the equilibrium switch at $\theta=\theta_L$ is socially optimal. Similarly, when $\theta=\theta_H$, we have $N_b=0$, $N_s=N$, $n=\frac{N_b}{N_s}=0$, so $\alpha(n)=\alpha(0)=0$. Therefore, $\Delta W(\theta_H,\lambda,n)=0$. Together with the arguments preceding (A.7) regarding why no single bank, fully internalizing all social benefits and costs from trading, finds it profitable to deviate from the benchmark no-screening equilibrium when $\theta\geq\theta_H$, this proves that the equilibrium switch at $\theta=\theta_H$ is socially optimal.

To prove the second part, we first show $\Delta W(\theta,\lambda,n)>0$ $\forall\theta\in(\theta_L,\theta_H)$, with n being determined by (15). Apply the implicit function theorem to (15): $\frac{\partial n}{\partial\lambda}=-\frac{1}{\lambda}\frac{\frac{1-\theta}{2}\frac{n-1}{n+1}-\frac{\psi}{(n+1)^2}}{\frac{1-\theta}{(n+1)^2}+\frac{2\psi}{(n+1)^3}}=-\frac{\frac{\Delta\pi^{\text{hom}}(\theta)}{\lambda^2(X-I)}}{\frac{1-\theta}{(n+1)^2}+\frac{2\psi}{(n+1)^3}}$, which is negative for $\theta\in(\theta_L,\theta^*)$ and positive for $\theta\in[\theta^*,\theta_H)$. For $\theta\in(\theta_L,\theta^*)$, rewrite the first expression for $\Delta W(\theta,\lambda,n)$ in (17) by substituting $\alpha(n)=\frac{\lambda n}{1+n}$, $p=X-\frac{X-I}{1+n}$, $N_b=N\frac{n}{1+n}$, $N_s=N\frac{1}{1+n}$ and replacing $\Delta\pi^{\text{hom}}(\theta)$ with the expression in (15):

$$\Delta W(\theta,\lambda,n)=[0.5(1-\theta)+\psi]N\frac{\lambda(X-I)}{(1+n)^2}>0. \qquad \text{(A.11)}$$

Because $\frac{\partial n}{\partial\lambda}<0$ and $\frac{\partial\Delta W(\theta,\lambda,n)}{\partial n}<0$ here, we have $\frac{d\Delta W(\theta,\lambda,n)}{d\lambda}=\frac{\partial\Delta W(\theta,\lambda,n)}{\partial n}\frac{\partial n}{\partial\lambda}+\frac{\partial\Delta W(\theta,\lambda,n)}{\partial\lambda}>0$.

Similarly, for $\theta\in[\theta^*,\theta_H)$ we can rewrite the second expression for $\Delta W(\theta,\lambda,n)$ in (17) as

$$\Delta W(\theta,\lambda,n)=0.5N(1-\theta)n^2\frac{\lambda(X-I)}{(1+n)^2}>0. \qquad \text{(A.12)}$$

Because $\frac{\partial n}{\partial\lambda}>0$ and $\frac{\partial\Delta W(\theta,\lambda,n)}{\partial n}>0$ here, we have $\frac{d\Delta W(\theta,\lambda,n)}{d\lambda}=\frac{\partial\Delta W(\theta,\lambda,n)}{\partial n}\frac{\partial n}{\partial\lambda}+\frac{\partial\Delta W(\theta,\lambda,n)}{\partial\lambda}>0$.

Then, we show n determined by (15) is not socially optimal. Consider $\theta\in(\theta_L,\theta^*)$ first. Substituting $\alpha(n)=\frac{\lambda n}{1+n}$, $p=X-\frac{X-I}{1+n}$, $N_b=N\frac{n}{1+n}$ and $N_s=N\frac{1}{1+n}$ into the first expression for $\Delta W(\theta,\lambda,n)$ in (17), we can rewrite it as

$$\Delta W(\theta,\lambda,n)$$
$$=N\left[\frac{\lambda n}{(1+n)^2}\frac{(1-\theta)(X-I)}{2}+\frac{\lambda n}{(1+n)^3}\psi(X-I)-\frac{1}{1+n}\Delta\pi^{\text{hom}}(\theta)\right]. \qquad \text{(A.13)}$$

Note that $\frac{\partial\Delta W(\theta,\lambda,n)}{\partial n}=N\Big[\frac{\lambda(1-n)}{(1+n)^3}\frac{(1-\theta)(X-I)}{2}+\frac{\lambda(1-2n)}{(1+n)^4}\psi(X-I)+\frac{1}{(1+n)^2}\Delta\pi^{\text{hom}}(\theta)\Big]$, which is decreasing in n, so the socially optimal n is given by the first-order condition (FOC), $\frac{\partial\Delta W(\theta,\lambda,n)}{\partial n}=0$:

$$\lambda(X-I)\left[\frac{1-\theta}{2}\frac{n-1}{n+1}-\psi\frac{1-2n}{(n+1)^2}\right]=\Delta\pi^{\text{hom}}(\theta). \qquad \text{(A.14)}$$

Compare (A.14) with the equilibrium condition under the private optimum (15): for a given $\theta\in(\theta_L,\theta^*)$ (so the RHSs equal in both conditions), the LHS in (A.14) is larger than the LHS in (15). Because the LHSs in both conditions are monotonically increasing in n, the socially optimal solution for n in (A.14) must be smaller than that in (15).

Similarly, for $\theta\in[\theta^*,\theta_H)$ we can rewrite the second expression for $\Delta W(\theta,\lambda,n)$ in (17) as

$$\Delta W(\theta,\lambda,n)$$
$$=N\left[\frac{\lambda n}{(1+n)^2}\frac{(1-\theta)(X-I)}{2}+\frac{\lambda n}{(1+n)^3}\psi(X-I)+\frac{n}{1+n}\Delta\pi^{\text{hom}}(\theta)\right]. \qquad \text{(A.15)}$$

With straightforward algebra, we can show that the socially optimal n here is also given by the same FOC in (A.14). Therefore, the socially optimal n is also smaller than the private optimum for every $\theta\in[\theta^*,\theta_H)$. □

Appendix B. Technical Details of Section 4 on Robustness and Extension

B.1. Robustness: Information Reusability

The buyer incurs a cost κc, with $\kappa>0$, at $t=1$ to screen the seller's loan. To compensate for that extra cost, the price p must be lowered (relative to the case with $\kappa=0$) to yield the buyer more surplus from trade[33]:

$$0.5(X-p)\geq\kappa c. \qquad \text{(B.1)}$$

Thus, $p\leq X-2\kappa c$. From (13), we know that the market tightness n must be reduced relative to the case with $\kappa=0$. The lowered n and p reduce a seller's gain from trade: (i) it becomes harder for a seller to meet a buyer with a lower n; and (ii) conditional on trading, the seller gains less with a lower p. Together, these may prevent a seller from emerging in the first place, thereby eliminating the heterogeneous equilibrium. We now formally show this.

Proof of Proposition 4. If $\kappa\geq\frac{X-I}{2c}$, then $p\leq I$; consequently, a seller gains nothing from trade because the loan sale proceeds are insufficient to finance the new asset. This eliminates the heterogeneous equilibrium for all θ values. Consider $\kappa<\frac{X-I}{2c}$, so $p\in(I,X-2\kappa c]$ and sellers may gain from trade. We show that a heterogeneous equilibrium still arises, albeit in a narrower range of θ values, as long as κ is not too large. We know from (13) and (B.1) that $p=X-\frac{X-I}{1+n}\leq X-2\kappa c$, so $n\leq\max\{\frac{X-I}{2\kappa c}-1,0\}=\frac{X-I}{2\kappa c}-1$ is needed for trading. For $\theta\in(\theta_L,\theta^*)$, the indifference condition for the heterogeneous equilibrium is

$$-\Delta\pi^{\text{hom}}(\theta)+0.5(1-\theta)\alpha(n)(p-I)$$
$$=(1-\theta)\frac{\alpha(n)}{n}[0.5(X-p)-\kappa c]+\psi\frac{\alpha(n)}{n}(X-p-2\kappa c)$$
$$=0.5(1-\theta)\frac{\alpha(n)}{n}(X-p)+\psi\frac{\alpha(n)}{n}(X-p)-(1-\theta+2\psi)\frac{\alpha(n)}{n}\kappa c, \qquad \text{(B.2)}$$

where a buyer's expected gain from trade conditional on meeting a seller is $0.5(X-p)-\kappa c$, and thus, the possibility of trading reduces the buyer's equity cost by $\psi\frac{\alpha(n)}{n}(X-p-2\kappa c)$. Comparing (B.2) with (A.5) (where $\kappa=0$), there is an extra term on the RHS of (B.2), $-(1-\theta+2\psi)\frac{\alpha(n)}{n}\kappa c$. For the same n, the RHS of (B.2) is thus smaller than the RHS of (A.5); the difference is larger with a bigger κ. Because the LHS is monotonically increasing in n, n determined by (B.2) must be smaller than that determined by (A.5), and that difference becomes larger when κ increases.

We can determine the cutoff value of κ above which the heterogeneous equilibrium does not arise. As long as $n < \frac{X-I}{2\kappa c} - 1$, we have $0.5(X-p) - \kappa c > 0$, so the RHS of (B.2) (a buyer's expected gain from trade) is positive. If at $n = \frac{X-I}{2\kappa c} - 1$, the LHS of (B.2) is negative, then the expected gain from trade is negative for a seller, so the heterogeneous equilibrium does not arise. Suppose at $n = \frac{X-I}{2\kappa c} - 1$, the LHS of (B.2) is positive. As κ increases, n decreases (as shown previously), so the LHS also falls. There exists a cutoff $\hat{\kappa}$, such that the LHS, which can be explicitly written as $-\Delta\pi^{\text{hom}}(\theta) + 0.5\lambda(1-\theta)(X-I)\left(\frac{n}{1+n}\right)^2$, turns negative for $\kappa > \hat{\kappa}$, eliminating the heterogeneous equilibrium. Formally, the cutoff $\hat{\kappa}$ is uniquely determined by (where $\hat{n}$ is the solution to (B.2) with $\kappa = \hat{\kappa}$)

$$0.5\lambda(1-\theta)(X-I)\left(\frac{\hat{n}}{1+\hat{n}}\right)^2 = \Delta\pi^{\text{hom}}(\theta). \tag{B.3}$$

Next, consider $\theta \in [\theta^*, \theta_H)$. The indifference condition for the heterogeneous equilibrium is

$$\begin{aligned}
&0.5(1-\theta)\alpha(n)(p-I)\\
&= \Delta\pi^{\text{hom}}(\theta) + (1-\theta)\frac{\alpha(n)}{n}[0.5(X-p) - \kappa c] + \psi\frac{\alpha(n)}{n}(X-p-2\kappa c)\\
&= \Delta\pi^{\text{hom}}(\theta) + 0.5(1-\theta)\frac{\alpha(n)}{n}(X-p) + \psi\frac{\alpha(n)}{n}(X-p)\\
&\quad - (1-\theta+2\psi)\frac{\alpha(n)}{n}\kappa c\\
&= \Delta\pi^{\text{hom}}(\theta) + \lambda[0.5(1-\theta)+\psi]\frac{X-I}{(1+n)^2} - \lambda(1-\theta+2\psi)\frac{\kappa c}{1+n}.
\end{aligned} \tag{B.4}$$

By the same argument as for $\theta \in (\theta_L, \theta^*)$, with $\theta \in [\theta^*, \theta_H)$, n determined by (B.4) must be smaller than that determined by the corresponding equilibrium condition (A.8) with $\kappa = 0$, and the difference (in term of n) becomes larger with a bigger κ. The LHS of (B.4) is positive as long as $p > I$, that is, $n > 0$. Note that $\Delta\pi^{\text{hom}}(\theta) < 0$ for $\theta \in [\theta^*, \theta_H)$, so we need $X - p - 2\kappa c = \frac{X-I}{1+n} - 2\kappa c$ to be sufficiently positive to keep the RHS also positive. As κ increases, n decreases (as shown previously), so the LHS also falls, hence the RHS also falls. Thus, the cutoff value of κ, $\tilde{\kappa}$, above which the heterogeneous equilibrium does not arise for $\theta \in [\theta^*, \theta_H)$, is uniquely determined by

$$[0.5(1-\theta)+\psi]\frac{X-I}{(1+\tilde{n})^2} - (1-\theta+2\psi)\frac{\tilde{\kappa}c}{1+\tilde{n}} = -\lambda^{-1}\Delta\pi^{\text{hom}}(\theta), \tag{B.5}$$

where $\tilde{n}$ is the solution to (B.4) with $\kappa = \tilde{\kappa}$. □

B.2. Extension: Capital Requirements

In the interest of not complicating the analysis, we assume that a minimum capital requirement $E^{\min}$ exists because of some social benefits of capital requirements outside the model. To make the point, we consider the case in which $E^{\min}$ is lower than the buyer's privately optimal capital, $E^{\text{het}}(\theta)$. This capital requirement is thus binding only for sellers who hold exactly $E^{\min}$.[34]

We first show that, interestingly, *even though capital requirements are not binding for buyers, a higher $E^{\min}$ indirectly causes each buyer to hold more equity* (*higher $E^{\text{het}}(\theta)$*). To see this, note that for a seller, holding $E^{\min}$ imposes a loss as it forces the bank to hold capital above its private optimum (which is zero). A positive $E^{\min}$ thus reduces the mass of sellers, relative to a regime of no capital requirements. The resulting increase in the market tightness (buyer/seller ratio) n leads to a higher loan price p. This has two adverse consequences for buyers. First, a higher n lowers the probability for a buyer to meet a seller, and a higher p reduces the buyer's gain from trade. Second, a larger n dilutes the incentive effect of trading, thereby reducing the equity-reduction advantage of trading enjoyed by the buyer (i.e., $E^{\text{het}}(\theta)$ rises; Lemma 1). The arguments show that capital requirements lower profits for both buyers and sellers and hence may eliminate the heterogeneous equilibrium. We now formally prove Proposition 5.

Proof of Proposition 5. Relative to the case with no capital requirement, with $E^{\min}$ a bank's expected profit without screening decreases by $\psi E^{\min}$, so the value-added of screening relative to no screening in a benchmark (no trading) equilibrium is now given by $\Delta\pi^{\text{hom}}(\theta) + \psi E^{\min}$, where $\Delta\pi^{\text{hom}}(\theta)$ is given by (11). Similar as (A.4) and (A.7), we can show that the cutoffs θ_L and θ_H are now determined by $-\Delta\pi^{\text{hom}}(\theta_L) - \psi E^{\min} + 0.5\lambda(1-\theta_L)(X-I) = 0$ and $\Delta\pi^{\text{hom}}(\theta_H) + \psi E^{\min} + 0.5\lambda(1-\theta_H)(X-I) + \psi\lambda(X-I) = 0$, respectively.

Apply the implicit function theorem to both previous conditions: $\frac{\partial\theta_L}{\partial E^{\min}} = \frac{\psi}{-\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta}|_{\theta=\theta_L} - 0.5\lambda(X-I)} > 0$ and $\frac{\partial\theta_H}{\partial E^{\min}} = \frac{\psi}{-\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta}|_{\theta=\theta_H} + 0.5\lambda(X-I)} > 0$.[35] We show $\frac{\partial\theta_L}{\partial E^{\min}} > \frac{\partial\theta_H}{\partial E^{\min}}$, so the region (θ_L, θ_H) with the heterogeneous equilibrium shrinks when $E^{\min}$ increases, that is, $\frac{\partial(\theta_H-\theta_L)}{\partial E^{\min}} < 0$, and a high enough $E^{\min}$ eliminates the region (θ_L, θ_H). To have $\frac{\partial\theta_L}{\partial E^{\min}} > \frac{\partial\theta_H}{\partial E^{\min}}$, it is sufficient to require $\frac{d^2\Delta\pi^{\text{hom}}(\theta)}{d\theta^2} = [\psi + 0.5(1-\theta)]\Upsilon''(\theta) - \Upsilon'(\theta) - \frac{4\psi(X-I+c)}{(1-\theta)^3} < 0$, so $-\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta}|_{\theta=\theta_H} > -\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta}|_{\theta=\theta_L}$. This clearly holds as $\Upsilon''(\theta) < 0$ and $\Upsilon'(\theta) > 0$. Finally, when λ increases, $\frac{\partial\theta_H}{\partial E^{\min}}$ decreases, whereas $\frac{\partial\theta_L}{\partial E^{\min}}$ increases, so $\frac{\partial(\theta_H-\theta_L)}{\partial E^{\min}}$ decreases, that is, $\frac{\partial^2(\theta_H-\theta_L)}{\partial E^{\min}\partial\lambda} < 0$: a larger λ increases the marginal impact of $E^{\min}$ in shrinking the region (θ_L, θ_H). □

Endnotes

[1] See Chan et al. (1986) and Di and Pattison (2020) for discussions of cross-sectional information reusability and its role in bank specialization in lending to borrowers in specific industries. Farhi and Tirole (2012) develop a theory in which banks make correlated asset portfolio choices, so banks in the cross section invest in similar loans.

[2] Our paper highlights one benefit of secondary market trading in terms of its ex ante effects. There are other benefits of trading that have been pointed out in previous research. For example, Faure-Grimaud and Gromb (2004) show that trading improves price informativeness and thereby increases a large shareholder's incentive to engage in value-enhancing activities, because it is more likely that the shareholder will be able to sell in the future at a price that reflects the value enhancement. In Aghion et al. (2004), greater price informativeness leads to better monitoring incentives inside the

firm for a similar reason. These papers essentially make the point that trading-induced enhancement in price informativeness improves exit options for a monitor/large shareholder and thus strengthens value-enhancing incentives inside firms. Our result about the benefit of trading complements this insight but differs in significant ways. First, in our model trading does not improve price informativeness per se but enables loan buyers to increase the marginal benefit of screening because of information reusability. Second, although trading provides loan sellers with exit options, the incentive effect in our model works through the *entry option* that trading creates in permitting high-capital banks to enter the secondary market and purchase loans. Third, and perhaps most significantly, a key point of our analysis is that the possibility of trading enables the emergence of an equilibrium with *endogenous* heterogeneity among ex ante identical banks and that welfare improves due to this heterogeneity. This result is not encountered in any of those previous studies on the benefits of trading.

[3] For the more recent literature on fire sales, with variations of the initial setting, see, for example, Carletti and Leonello (2019), Dow and Han (2018), Kurlat (2016), and Lorenzoni (2008).

[4] For a bank, the decision to make a loan is similar to a nonfinancial firm's capital budgeting decision of whether to invest in a project based on pre-investment screening.

[5] This equal-prior assumption, made for algebraic simplicity, does not qualitatively affect our results.

[6] Assuming that screening yields a perfect signal is a simplification that greatly facilitates analytical tractability and allows us to explicitly examine equilibrium properties and conduct welfare analysis. If screening precision is exogenously fixed, then the analysis remains largely unaffected even if the signal is informative but noisy. However, allowing the signal to be noisy with its precision being endogenously chosen from a continuum $[0.5, 1]$ leads to a much more complex model with the equilibrium being characterized with a highly nonlinear system of equations, which impedes tractability.

[7] Screening would have been redundant if the bank were to approve a loan revealed to be M.

[8] A smaller κ means information is more reusable. Information is fully reusable if $\kappa = 0$.

[9] See, for example, Beck et al. (2018), Bolton et al. (2016b), Boot (2000), Boot et al. (1993), and Sharpe (1990).

[10] Bharath et al. (2007, p. 370) write: "To the extent that relationship lending produces reusable and proprietary information about the borrower, a possible benefit for the relationship lender is that it would be better placed to win future loan business and other fee-generating services from its relationship borrower." They document that the probability of providing a future loan is 42% for a relationship lender and 3% for a nonrelationship lender.

[11] In our model, information reusability leads to a lower information processing cost for the second loan than for the first. If one were to assume that the benefit of information reusability is unbounded, then the bank would be a natural monopolist in the first best. Some theories have introduced features that lead to a finite optimal bank size (Krasa and Villamil 1992). Others have discussed that if banks were allowed to grow without restraint, there would be additional social costs arising from the excessive concentration of systemic risk in a few institutions, governance challenges, and too-big-to-fail subsidies (Laeven et al. 2014). We take a reduced-form approach and assume that capacity constraints prevent the bank from screening more than two loans.

[12] All results are qualitatively unaffected by the linearity specification of the cost, which is made for algebraic brevity.

[13] An M loan's date 2 payoff is X in a boom $(s = b)$, but there will be no trading in a boom, as shown later.

[14] Purchasing an M loan whose date 2 payoff is zero in a recession would cause the buyer to repay the new debt out of the date 2 cash flows of its own legacy loan and new asset, which is never optimal. That is, although the loan purchase is funded with outside finance, the buyer is not fully protected by limited liability. This holds even if the M loan's date 2 payoff is random in a recession, say X (with probability q) or zero (with probability $1 - q$), instead of zero with certainty. To see this, first note $p > qX$ (the seller would not sell its loan if $p \leq qX$). The question is whether the buyer, with outside finance, would be willing to purchase M by paying more than the loan's expected value (qX). The buyer's expected net payoff from purchasing M would be $q(X - p) - (1 - q)\min\{p, (X - D) + (X - I)\}$. To understand this, note that the remaining cash flows from the buyer's own legacy loan (which is G and pays X) and the new asset (which pays X), after paying off its own legacy debt (D) and new debt (I) raised to finance the new asset, is $(X - D) + (X - I)$. Therefore, when the M loan returns zero (with probability $1 - q$), the buyer repays $\min\{p, (X - D) + (X - I)\}$ to financiers who provided p to fund the loan purchase: the buyer is again not fully protected by limited liability. If $p \leq (X - D) + (X - I)$, the net payoff is $qX - p < 0$. If $p > (X - D) + (X - I)$, the net payoff is $q(X - p) - (1 - q)[(X - D) + (X - I)] < qX - 2(X - I)$, which is negative if $q < 2\left(1 - \frac{I}{X}\right)$. Meanwhile, we need $[\overline{\theta} + (1 - \overline{\theta})q]X < I$ (i.e., an M loan has a negative date 0 NPV), which leads to $q < \left(\frac{I}{X} - \overline{\theta}\right)(1 - \overline{\theta})^{-1}$. Therefore, a sufficient condition to ensure a negative net payoff from buying an M loan is $\left(\frac{I}{X} - \overline{\theta}\right)(1 - \overline{\theta})^{-1} \leq 2\left(1 - \frac{I}{X}\right)$, that is, $\frac{I}{X} \leq \frac{2 - \overline{\theta}}{3 - 2\overline{\theta}}$, which holds if $\overline{\theta}$, albeit smaller than $\frac{I}{X}$, is sufficiently close to $\frac{I}{X}$.

[15] Recall that the rest of the financing for the new asset, δI, comes from pledging the new asset's own date 2 cash flow.

[16] The reliance on $Y(\theta)$ for this result is necessitated by the model's simplifying stark specification that screening is either perfect or uninformative. If the screening precision were (endogenously) chosen from a continuum, then the equilibrium equity would be decreasing in θ even absent $Y(\theta)$. However, the discrete specification of precision greatly aids tractability.

[17] This monotonicity ensures that any characterized equilibrium (screening or no screening; homogeneous or heterogeneous) obtains in connected (but not disjoint) regions of θ; see Propositions 1 and 2. This facilitates the interpretation of our results.

[18] The proof of Proposition 1 in Appendix A shows that there exists $E' > E^{\text{hom}}(\theta)$, such that for any $E \in [0, E')$ the term $-E - \psi E + [\theta + 0.5(1 - \theta)](X - D + X - I) - 0.5(1 - \theta)Y(\theta)$ is strictly positive (hence, exceeds $-\psi E$), given (4). The solution to (7) is thus given by (8). With $E = E^{\text{hom}}(\theta)$ in equilibrium, if the bank were to avoid screening, it would approve (rather than reject) the unscreened loan. This case is economically meaningful. If a bank that chooses not to screen despite positive equity were to reject the unscreened loan, it would have been strictly better off by not holding any (costly) equity in the first place, avoiding screening and unconditionally rejecting loans. The only benefit of avoiding screening with positive equity is to make loans that can be funded with cheap debt because investors mistakenly believe the bank has screened. Moreover, that the RHS of (7) being strictly positive in equilibrium also implies the bank's expected net profit from screening (LHS) is strictly positive, so screening is economically viable.

[19] This is the standard skin in the game argument. Both bank equity and charter value provide skin in the game, so it is useful to note their difference. As θ increases, the bank's screening incentive is affected in two ways. First, an M loan returns X in a boom, making the earlier screening a waste, so a higher θ weakens the screening incentive, and more equity is needed to restore this incentive. Second, the charter value is increasing in θ, so a higher θ increases the loss of charter value upon default. The condition in (2) ensures that

the second effect dominates, so an increase in θ leads to the bank decreasing its reliance on equity, $\frac{dE^{\text{hom}}(\theta)}{d\theta} < 0$. The condition in (3) ensures that the charter value by itself is insufficient to provide the bank with all the incentive needed to screen, so some equity is warranted. Specifically, having $\frac{dE^{\text{hom}}(\theta)}{d\theta} < 0$ requires $Y'(\theta) > \frac{2(X-I+c)}{(1-\theta)^2}$, which holds for any $\theta \in [0, \overline{\theta}]$ given $Y''(\theta) < 0$ and (2).

[20] Holding $E \in (0, E^{\text{hom}}(\theta))$ imposes on the bank the cost of that equity but financiers still believe the bank does not screen.

[21] At $\theta = \theta^*$, $\pi_{\text{s}}^{\text{hom}}(\theta^*) = \pi_{\text{ns}}^{\text{hom}}(\theta^*)$, and banks do not screen; this is innocuous. Assuming intermediate values of c precludes the uninteresting and unrealistic outcome that the same type of equilibrium (screening or no-screening) prevails for all θ values.

[22] In this case, the loan is equally likely to be G or M as perceived by the buyer, so its expected date 2 cash flow is $0.5X$. The buyer would not purchase this loan, given $p \geq I > 0.5X$, where the first inequality follows from (13) and the second inequality follows from (1). The buyer's expected net payoff from buying this loan would be $0.5(X-p) - 0.5\min\{p, 0.5(X-I) + (X-I)\}$, where the remaining cash flow from the buyer's own (unscreened) legacy loan after repaying its legacy debt (I) is $0.5(X-I)$, and the remaining cash flow from the buyer's new asset after repaying the new debt (I) raised to fund that asset is $X - I$ (these are used to repay financiers who provided p to fund the loan purchase when the purchased loan returns zero). If $p \leq 1.5(X-I)$, the net payoff is $0.5X - p < 0$. If $p > 1.5(X-I)$, the net payoff is $0.5(X-p) - 0.75(X-I) < 0.5[(X-p) - (X-I)] = 0.5(I-p) \leq 0$. This also verifies the presumption that buyers raise equity and screen (while sellers do not).

[23] The probability of a recession, $1-\theta$, does not appear in the equity comparison in (14). This is because the terms in (14) correspond to what the bank would lose *conditional on a recession* in the absence of screening. The probability of a recession therefore does not affect these terms.

[24] Let $n = 1$ in (15), we have $-0.25\lambda\psi(X-I) = \Delta\pi^{\text{hom}}(\theta^{\max})$. Note that $\theta^{\max} > \theta^*$, because $\Delta\pi^{\text{hom}}(\theta^{\max}) < 0$, $\Delta\pi^{\text{hom}}(\theta^*) = 0$, and $\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta} < 0$.

[25] The two expressions for $\Delta W(\theta, \lambda, n)$ in (17) coincide at $\theta = \theta^*$ because $\Delta\pi^{\text{hom}}(\theta^*) = 0$.

[26] A higher λ also expands the region (θ_L, θ_H) in which the heterogeneous equilibrium arises (Proposition 2). Therefore, a more efficient interbank market not only increases the occurrence of a heterogeneous equilibrium but also leads to a larger welfare enhancement (relative to the benchmark equilibrium) conditional on having the heterogeneous equilibrium.

[27] The usual justification for a capital requirement is that, without it, banks would keep too low a capital ratio that hurts the public safety net (Merton 1977). As we show in Proposition 5, if a capital requirement (designed to protect the public safety net) is binding, then it can reduce heterogeneity, which is welfare depleting. Absent a safety-net reason for a capital requirement, the bank's capital choice is to provide the appropriate screening incentive, and banks make this choice optimally in the screening equilibrium. The divergence of the private optimum from the social optimum in our model is that there are too many high-capital banks in the heterogeneous equilibrium, and each keeps higher capital than is socially optimal (Proposition 3). This is not an inefficiency that capital requirements can address.

[28] Indeed, this was a stated goal of the Basel I Capital Accord. Of course, despite capital requirements other factors can create interbank heterogeneity in capital ratios, like politics (Thakor 2021) and the extent of institutional ownership in the bank (Garel et al. 2022). See Section 4.3 for further discussion.

[29] Recent empirical evidence indicates that the fire-sales discount may be smaller than previously estimated; see Franks et al. (2021).

[30] Such subsidies can take the form of the central bank providing liquidity in its lender-of-last-resort function (Alves et al. 2021).

[31] It can be verified that with $E = E^{\text{het}}(\theta)$ in (14), $-E - \psi E + [\theta + 0.5(1-\theta)](X - D + X - I) - 0.5(1-\theta)Y(\theta) > -\psi E$, which follows from (4). Therefore, the RHS of (A.2) equals $-E - \psi E + [\theta + 0.5(1-\theta)](X - D + X - I) - 0.5(1-\theta)Y(\theta)$ for any $E \in [0, E'')$, with some $E'' > E^{\text{het}}(\theta)$, based on which $E \geq E^{\text{het}}(\theta)$ obtains.

[32] The RHS uses the fact that a buyer will not trade in the interbank market without screening; see Endnote 22.

[33] The buyer will not purchase a loan without screening it (see Endnote 22). Before screening, the buyer expects to buy the seller's loan with probability 0.5 (as the seller's unscreened legacy loan is equally likely to be G or M in a recession), and gain $X - p$ from the purchase. The LHS of (B.1) is thus the buyer's expected gain from trade, which must be no lower than the cost of screening before trade, κc, on the RHS.

[34] If $E^{\min}$ is so high that it is binding for all banks, then it is trivially true that the heterogeneous equilibrium vanishes.

[35] Note that $\frac{d\Delta\pi^{\text{hom}}(\theta)}{d\theta} < -0.5\lambda(X-I)$, ensured by the condition in (5).

References

Acharya VV, Skeie D (2011) A model of liquidity hoarding and term premia in inter-bank markets. *J. Monetary Econom.* 58(5):436–447.

Acharya VV, Yorulmazer T (2008) Cash-in-the-market pricing and optimal resolution of bank failures. *Rev. Financial Stud.* 21(6):2705–2742.

Acharya VV, Gromb D, Yorulmazer T (2012) Imperfect competition in the interbank market for liquidity as a rationale for central banking. *Amer. Econom. J. Macroeconom.* 4(2):184–217.

Acharya VV, Shin HS, Yorulmazer T (2013) A theory of arbitrage capital. *Rev. Corporate Finance Stud.* 2(1):62–97.

Aghion P, Bolton P, Tirole J (2004) Exit options in corporate finance: Liquidity vs. incentives. *Rev. Finance* 8(3):327–353.

Allen F, Gale D (1994) Limited market participation and volatility of asset prices. *Amer. Econom. Rev.* 84(4):933–955.

Allen F, Carletti E, Marquez R (2011) Credit market competition and capital regulation. *Rev. Financial Stud.* 24(4):983–1018.

Allen F, Carletti E, Marquez R (2015) Deposits and bank capital structure. *J. Financial Econom.* 118(3):601–619.

Alves N, Bonfim D, Soares C (2021) Surviving the perfect storm: The role of the lender of last resort. *J. Financial Intermediation* 47:100918.

Beck T, Degryse H, De Haas R, Van Horen N (2018) When arm's length is too far: Relationship banking over the credit cycle. *J. Financial Econom.* 127(1):174–196.

Begenau J, Bigio S, Majerovitz J, Vieyra M (2020) A q-theory of banks. Technical report, National Bureau of Economic Research, Cambridge, MA.

Berger AN, Bouwman CH (2013) How does capital affect bank performance during financial crises? *J. Financial Econom.* 109(1):146–176.

Bharath S, Dahiya S, Saunders A, Srinivasan A (2007) So what do I get? The bank's view of lending relationships. *J. Financial Econom.* 85(2):368–419.

Bharath S, Dahiya S, Saunders A, Srinivasan A (2011) Lending relationships and loan contract terms. *Rev. Financial Stud.* 24(4):1141–1203.

Bhat G, Desai HA (2020) Bank capital and loan monitoring. *Accounting Rev.* 95(3):85–114.

Bolton P, Santos T, Scheinkman JA (2016a) Cream-skimming in financial markets. *J. Finance* 71(2):709–736.

Bolton P, Freixas X, Gambacorta L, Mistrulli PE (2016b) Relationship and transaction lending in a crisis. *Rev. Financial Stud.* 29(10):2643–2676.

Boot AW (2000) Relationship banking: What do we know? *J. Financial Intermediation* 9(1):7–25.

Boot AW, Greenbaum SI, Thakor AV (1993) Reputation and discretion in financial contracting. *Amer. Econom. Rev.* 83(5):1165–1183.

Bord VM, Ivashina V, Taliaferro RD (2021) Large banks and small firm lending. *J. Financial Intermediation* 48:100924.

Bostandzic D, Weiß GN (2018) Why do some banks contribute more to global systemic risk? *J. Financial Intermediation* 35:17–40.

Bouwman CH (2019) Creation and regulation of bank liquidity. Berger AN, Molyneux P, Wilson JO, eds. *Oxford Handbook of Banking* (Oxford University Press, Oxford, UK), 181–228.

Bouwman CH, Kim H, Shin S (2018) Bank capital and bank stock performance. Technical report, Texas A&M University, College Station, TX.

Boyd JH, Prescott EC (1986) Financial intermediary-coalitions. *J. Econom. Theory* 38(2):211–232.

Boyson N, Helwege J, Jindra J (2014) Crises, liquidity shocks, and fire sales at commercial banks. *Financial Management* 43(4):857–884.

Camargo B, Lester B (2014) Trading dynamics in decentralized markets with adverse selection. *J. Econom. Theory* 153:534–568.

Carletti E, Leonello A (2019) Credit market competition and liquidity crises. *Rev. Finance* 23(5):855–892.

Carletti E, Marquez R, Petriconi S (2020) The redistributive effects of bank capital regulation. *J. Financial Econom.* 136(3):743–759.

Chan YS, Greenbaum SI, Thakor AV (1986) Information reusability, competition and bank asset quality. *J. Banking Finance* 10(2): 243–253.

Chiu J, Koeppl TV (2016) Trading dynamics with adverse selection and search: Market freeze, intervention and recovery. *Rev. Econom. Stud.* 83(3):969–1000.

Choi DB (2014) Heterogeneity and stability: Bolster the strong, not the weak. *Rev. Financial Stud.* 27(6):1830–1867.

Coval JD, Thakor AV (2005) Financial intermediation as a beliefs-bridge between optimists and pessimists. *J. Financial Econom.* 75(3):535–569.

Di W, Pattison N (2020) Distant lending, specialization, and access to credit. Working Paper, Federal Reserve Bank of Dallas, TX.

Dow J, Han J (2018) The paradox of financial fire sales: The role of arbitrage capital in determining liquidity. *J. Finance* 73(1): 229–274.

Eisenbach TM, Lucca DO, Townsend RM (2019) The economics of bank supervision. Report, Federal Reserve Bank of New York, NY.

Farhi E, Tirole J (2012) Collective moral hazard, maturity mismatch, and systemic bailouts. *Amer. Econom. Rev.* 102(1):60–93.

Faure-Grimaud A, Gromb D (2004) Public trading and private incentives. *Rev. Financial Stud.* 17(4):985–1014.

Franks JR, Seth G, Sussman O, Vig V (2021) Revisiting the asset fire sale discount: Evidence from commercial aircraft sales. Working paper, European Corporate Governance Institute, Brussels, Belgium.

Furlong FT, Kwan S (2006) Sources of bank charter value. Working paper, Federal Reserve Bank of San Francisco, CA.

Gale D, Yorulmazer T (2013) Liquidity hoarding. *Theoretical Econom.* 8(2):291–324.

Gambacorta L, Shin HS (2018) Why bank capital matters for monetary policy. *J. Financial Intermediation* 35:17–29.

Garel A, Petit-Romec A, Vander Vennet R (2022) Institutional shareholders and bank capital. *J. Financial Intermediation* 50:100960.

Goldstein I, Kopytov A, Shen L, Xiang H (2020) Bank heterogeneity and financial stability. Technical report, University of Pennsylvania, Philadelphia, PA.

Granja J, Matvos G, Seru A (2017) Selling failed banks. *J. Finance* 72(4):1723–1784.

Greenbaum SI, Thakor AV, Boot AW (2019) *Contemporary Financial Intermediation* (Academic Press).

Hirtle B, Kovner A, Plosser MC (2019) The impact of supervision on bank performance. Report, Federal Reserve Bank of New York, NY.

Holmstrom B, Tirole J (1997) Financial intermediation, loanable funds, and the real sector. *Quart. J. Econom.* 112(3):663–691.

Hosios AJ (1990) On the efficiency of matching and related models of search and unemployment. *Rev. Econom. Stud.* 57(2):279–298.

Irani RM, Meisenzahl RR (2017) Loan sales and bank liquidity management: Evidence from a US credit register. *Rev. Financial Stud.* 30(10):3455–3501.

Irani RM, Iyer R, Meisenzahl RR, Peydro J-L (2021) The rise of shadow banking: Evidence from capital regulation. *Rev. Financial Stud.* 34(5):2181–2235.

Jiang E, Matvos G, Piskorski T, Seru A (2020) Banking without deposits: Evidence from shadow bank call reports. Technical report, National Bureau of Economic Research, Cambridge, MA.

Jorge J, Kahn CM (2017) Liquidity freezes under adverse selection. ECB Working Paper, European Central Bank, Frankfurt, Germany.

Kiyotaki N, Wright R (1993) A search-theoretic approach to monetary economics. *Amer. Econom. Rev.* 83(1):63–77.

Kovner A, Van Tassel P (2019) Evaluating regulatory reform: Banks' cost of capital and lending. Report, Federal Reserve Bank of New York, NY.

Krasa S, Villamil AP (1992) A theory of optimal bank size. *Oxford Econom. Papers* 44(4):725–749.

Kurlat P (2016) Asset markets with heterogeneous information. *Econometrica* 84(1):33–85.

Laeven L, Ratnovski L, Tong H (2014) Bank size and systemic risk. IMF Working Paper, International Monetary Fund, Washington, D.C.

Lorenzoni G (2008) Inefficient credit booms. *Rev. Econom. Stud.* 75(3): 809–833.

Mehran H, Thakor A (2011) Bank capital and value in the cross-section. *Rev. Financial Stud.* 24(4):1019–1067.

Merton RC (1977) An analytic derivation of the cost of deposit insurance and loan guarantees: An application of modern option pricing theory. *J. Banking Finance* 1(1):3–11.

Myers SC, Majluf NS (1984) Corporate financing and investment decisions when firms have information that investors do not have. *J. Financial Econom.* 13(2):187–221.

Novo-Peteiro JA (2000) New technologies, information reusability and diversification: A simple model of a banking firm. *Inform. Econom. Policy* 12(1):69–88.

Paravisini D, Rappoport V, Schnabl P (2015) Specialization in bank lending: Evidence from exporting firms. Technical report, National Bureau of Economic Research, Cambridge, MA.

Pérignon C, Thesmar D, Vuillemey G (2018) Wholesale funding dry-ups. *J. Finance* 73(2):575–617.

Philippon T, Skreta V (2012) Optimal interventions in markets with adverse selection. *Amer. Econom. Rev.* 102(1):1–28.

Ramakrishnan RT, Thakor AV (1984) Information reliability and a theory of financial intermediation. *Rev. Econom. Stud.* 51(3):415–432.

Saunders A, Wilson B (2001) An analysis of bank charter value and its risk-constraining incentives. *J. Financial Service Res.* 19(2):185–195.

Sharpe SA (1990) Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *J. Finance* 45(4):1069–1087.

Shleifer A, Vishny R (2011) Fire sales in finance and macroeconomics. *J. Econom. Perspective* 25(1):29–48.

Shleifer A, Vishny RW (1992) Liquidation values and debt capacity: A market equilibrium approach. *J. Finance* 47(4):1343–1366.

Thakor AV (2014) Bank capital and financial stability: An economic trade-off or a faustian bargain? *Annu. Rev. Financial Econom.* 6(1):185–223.

Thakor AV (2021) Politics, credit allocation and bank capital requirements. *J. Financial Intermediation* 45:100820.

Tirole J (2012) Overcoming adverse selection: How public intervention can restore market functioning. *Amer. Econom. Rev.* 102(1): 29–59.