Bank Heterogeneity and Financial Stability*

Itay Goldstein,[†] Alexandr Kopytov,[‡] Lin Shen,[§] and Haotian Xiang[¶]

December 21, 2023

Abstract

We propose a model of the financial system in which banks are individually prone to runs and connected through fire sales. Strategic complementarities within and across banks amplify each other, making heterogeneity in bank risks a key factor shaping the fragility of each bank and the entire system. As long as different banks are interconnected, an increase in heterogeneity stabilizes all banks. Reductions in asset commonality, bank-specific disclosures, and even broad-based policies such as asset purchases and liquidity requirements can enhance stability by increasing bank heterogeneity.

JEL: G01, G21, G28

*We thank Philipp Schnabl (editor), an anonymous referee, Franklin Allen, Max Bruche, Markus Brunnermeier, Stephen Cecchetti, Jean Dermine, Thomas Eisenbach, Piero Gottardi, Naveen Khanna, John Kuong, Xuewen Liu, Zehao Liu, Andrey Malenko, Dmitry Orlov, Cecilia Parlatore, Adriano Rampini, Stathis Tompaidis, Aleh Tsyvinski, Harald Uhlig, Ming Yang, Yao Zeng, Zhen Zhou, as well as seminar participants at INSEAD, Princeton, UBC, Washington University in St. Louis, NYU, UCSD, Nankai, AFA, Virtual Finance Theory Seminar, NBER Summer Institute, RCFS/RAPS Winter Conference, International Moscow Finance Conference, Banca d'Italia and Bocconi University Conference, Scottish Economic Society Annual Conference and CICF for useful comments. Earlier versions of the paper circulated under the title "Synchronicity and Fragility."

[†]University of Pennsylvania and NBER, 2253 Steinberg-Dietrich Hall, 3620 Locust Walk, Philadelphia, PA 19104, itayg@wharton.upenn.edu (Corresponding author)

[‡]University of Rochester, 3-141 Carol Simon Hall, 500 Wilson Blvd, Rochester, NY 14627, akopytov@simon.rochester.edu

[§]INSEAD, Bd de Constance, 77300 Fontainebleau, France, lin.shen@insead.edu

[¶]Peking University, No.5 Yiheyuan Rd, Haidian Dist, Beijing 100871, xiang@gsm.pku.edu.cn

1 Introduction

The global financial crisis of 2007–2008 put the correlation in risks across financial institutions (for short, banks) in the limelight. The concern emphasized by policymakers (Haldane, 2009; Yellen, 2013) is that the joint failure of many banks presents a big threat to the economy. Therefore, several measures have been developed to track comovement in banks' risks.[1] What is less clear is whether the correlation between banks makes them more prone to failure overall. This is critical for evaluating the importance of tail comovement. After all, if banks were more correlated in their failures, but at the same time less likely to fail overall, then the implications of comovement for financial fragility would be much less severe. The model presented in this paper uncovers a new channel through which risk correlation makes banks overall more fragile. Based on this channel, we develop new implications for financial policy aimed at minimizing fragility.

Our model features banks that are individually fragile, due to the provision of liquidity transformation, and indirectly interconnected through fire-sale spillovers.[2] When facing withdrawals, banks liquidate their assets in a common market, thereby imposing negative fire-sale externalities on one another. We define financial fragility as the unconditional probability of individual banks to suffer a run. We show that the degree of bank heterogeneity—the extent to which different banks experience runs in different states and a measure of tail-risk dispersion in our setting—is a key indicator of financial stability of individual banks and the system as a whole. We further argue that boosting heterogeneity, at least to a certain point, is a natural goal for regulators aiming at minimizing real losses caused by financial fragility and inefficient fire sales.

Why does bank heterogeneity reduce fragility? The main economic force behind this result is a two-layered coordination problem that naturally arises in our setting. First, typical to a bank-run model, there is a *within-bank strategic complementarity*. Investors who withdraw money from a bank early cause costly asset liquidations and impose neg-

[1] Popular systemic risk measures are based on tail comovement among financial institutions, e.g. Huang, Zhou, and Zhu (2009), Adrian and Brunnermeier (2016), Acharya, Pedersen, Philippon, and Richardson (2017), Brownlees and Engle (2017).

[2] We refer to individual institutions as banks, but our analyses can be applied to other types of institutions with runnable liabilities (e.g., corporate bond mutual funds, Goldstein, Jiang, and Ng, 2017).

ative externalities on those who stay. Second, there is a *cross-bank strategic complementarity* due to fire-sale spillovers. Key to our results is that these two complementarities amplify each other. An investor is more concerned about withdrawals in her bank when she expects investors of other banks to withdraw in their banks, forcing premature liquidations and driving up the fire-sale discount. Therefore, if withdrawals from different banks coincide with each other, cross-bank fire-sale spillovers are particularly detrimental to bank stability. As withdrawals become more dispersed, fire-sale spillovers are attenuated and become less detrimental to stability. There exists a natural limit to this result: If runs are sufficiently dispersed, investors become certain about run situations in other banks. In that case, the strength of fire-sale spillovers is effectively fixed, and bank heterogeneity no longer affects fragility.

It is important to highlight that our theory offers very distinct insights compared to existing theories about systemic risk, which typically feature a tension between systemic risk and risks of individual institutions. In particular, as interconnectedness rises, banks are more likely to fail together but weaker banks are less likely to fail.[3] In our paper, such tension does not normally exist, since larger heterogeneity across banks makes each one of them less likely to fail in the first place. Hence, heterogeneity is a first-order concern. This highlights the need to track heterogeneity. Ignoring the impact of policies on heterogeneity would lead to an imprecise assessment of their effect on financial stability.

How can regulators affect bank heterogeneity? To answer this question, we identify main factors contributing to heterogeneity and point out how various policies affect it. First, if banks hold different assets, they are exposed to different fundamental shocks, which naturally leads to dispersion in withdrawals. Policies that reduce asset commonality, such as ring fencing, can thus increase heterogeneity. Second, because investors are the ones who make withdrawal decisions, what matters for runs is perceived rather than physical asset commonality. By mandating disclosure of bank-specific information, regulators can reduce investors' perceived asset commonality and boost heterogeneity. Third, since run decisions are shaped by investors' beliefs about the magnitude of the fire-sale dis-

[3]See, for example, Cabrales, Gottardi, and Vega-Redondo (2017) in the context of financial networks and Bouvard, Chaigneau, and Motta (2015) in the context of bank disclosure.

count, liquidity conditions in the asset market also play an important role. We show that secondary market liquidity injections and liquidity requirements have a stronger positive effect on liquidity conditions perceived by investors of banks with stronger fundamentals and thus increase heterogeneity.

We now describe the model and the results in more detail. As is typical for models of financial fragility, run decisions of bank investors are characterized by run thresholds. Specifically, run happens if and only if the aggregate fundamental falls below a bank-specific threshold. Therefore, low run thresholds indicate low fragility. In our model, all bank-specific thresholds are determined jointly in equilibrium. When deciding whether to withdraw their funds early, investors assess run situations in all banks in the financial system. To pin down equilibrium run thresholds, we thus need to characterize investors' beliefs about run decisions of other investors in their banks and investors in other banks. Analytical characterization of these beliefs in a setting featuring interacting within- and cross-bank complementarities is one contribution of our paper.

There are two main factors that shape run thresholds. The first one is bank fundamentals. In the model, bank asset returns are subject to aggregate and idiosyncratic shocks. Idiosyncratic shocks separate banks into two types ex post: strong banks receive a positive shock and weak banks receive a negative shock. Strong banks are relatively more stable, and a worse aggregate shock is needed to trigger runs on these banks. Hence, they have a lower run threshold. Second, run thresholds depend on investors' beliefs about the fire-sale discount. Marginal strong-bank investors expect a higher fire-sale discount than marginal weak-bank investors do. This is because if strong-bank investors are on the margin of running or not, they expect that weak banks are experiencing severe run problems. Conversely, marginal investors of weak banks expect fewer runs on strong banks. Overall, reflecting both the fundamental effect and the fire-sale effect, the run thresholds of both types of banks are determined in equilibrium. The distance between the run thresholds of strong and weak banks captures *bank heterogeneity*.

Our main theoretical result is that an increase in bank heterogeneity makes both weak and strong banks more resilient to panic runs by weakening the reinforcement between within- and cross-bank complementarities. A larger heterogeneity implies that strong banks' stability is challenged by a greater downward pressure that weak banks impose on liquidation prices, and weak banks' fragility is alleviated by a lower pressure that strong banks impose. The key to the decrease in the overall fragility is that the effect on weak banks dominates that on strong banks. This is a direct result of the fact that within- and cross-bank complementarities are mutually reinforcing. In particular, given that weak banks are more internally fragile, their investors are more strongly affected by the alleviated fire-sale pressure than strong-bank investors are affected by the intensified fire-sale pressure. Hence, fragility is lower when heterogeneity increases.

When heterogeneity becomes too large, however, a further increase in it has no effect on fragility. In this case, strong banks and weak banks are effectively disconnected. Specifically, investors in strong (weak) banks become certain that all (none) investors in weak (strong) banks will run. Therefore, when investors decide whether to run, they face uncertainties only regarding run situations in the same type of banks. As a result, an increase in heterogeneity no longer affects cross-bank spillovers. These extreme scenarios with two non-interacting groups of banks are less interesting from a theoretical standpoint. Furthermore, they are also difficult to reconcile with experiences from recent financial crises that featured substantial uncertainties about the health of various financial institutions among market participants.

After establishing the key insight about heterogeneity and fragility, we proceed to policy analyses. We show that in our model, reducing bank fragility is welfare-improving because runs are associated with real resource losses due to fire sales. Thus, stabilizing financial sector is a natural goal for the regulator. We consider four policies, often discussed as part of financial regulation, and study their effect on fragility through the heterogeneity channel.

First, we consider a ring-fencing policy that separates bank balance sheets into different divisions according to business or geographic focus. Such a policy increases the dispersion in asset returns across different bank divisions. In the model, this corresponds to an increase in the ex-post difference between weak and strong banks and, hence, a higher bank heterogeneity. As argued above, an increase in heterogeneity is beneficial for stability of all banks (unless heterogeneity is already too large). The existing literature that examines the relation between asset commonality and financial stability has mainly focused on cascades of fundamental defaults in which investors play a passive role. This literature suggests that asset differentiation reduces systemic bank failures at the expense of more individual failures of weak banks (Shaffer, 1994; Wagner, 2010 and 2011; Ibragimov, Jaffee, and Walden, 2011). Simulation analyses suggest that purely fundamental default cascades are unlikely in modern financial systems (Elsinger, Lehar, and Summer, 2006; Upper, 2011). Differently, we consider panic-based runs with investors actively withdrawing because of concerns about others' withdrawals, precipitating illiquidity which then spreads across banks through fire-sale spillovers. Importantly, we show that in our environment asset differentiation can make both strong and weak banks more resilient to panic-driven runs.

Second, we consider regulatory disclosure that affects the quality of bank-specific information available to investors. We extend the model by adding noise to investors' information about bank-specific shocks. In this extended setting, we show that what matters for heterogeneity is perceived differences between bank asset returns. In an opaque financial system, investors can hardly distinguish between strong and weak banks, resulting in an effectively homogeneous financial system. Disclosing bank-specific information enlarges heterogeneity and stabilizes the financial system.[4] While existing literature argues that disclosing bank-specific information can undermine the stability of weak banks (e.g., Bouvard et al., 2015 and Goldstein and Leitner, 2018), our results suggest that disclosure can stabilize the entire financial sector, including weak banks, by alleviating fire-sale pressure on particularly fragile institutions.[5]

The third policy we examine is secondary market liquidity injections, widely used by regulators during market turmoils in the 2007–2008 financial crisis and the COVID-

[4] A contemporaneous paper by Dai, Luo, and Yang (2021) argues that disclosure of banks' systemic risk exposures—but not their idiosyncratic risks—can mitigate financial fragility. In our framework with reinforcing complementarities, disclosing bank-specific information can enlarge heterogeneity which is beneficial for overall stability.

[5] In a one-bank setting, Parlatore (2015) shows that an increase in transparency can worsen coordination problem among depositors and thus undermine stability.

19 pandemic. Unlike ring-fencing and disclosure policies, a liquidity injection does not affect (perceived) dispersion in bank fundamentals. However, it does affect heterogeneity by reshaping investors' beliefs about the fire-sale discount. In the model, a liquidity injection reduces the fire-sale discount for any amount of long-term assets liquidated by banks. Surprisingly, even though an injection is broad-based—that is, the regulator does not purchase assets owned by a particular group of banks—it tends to reduce the fire-sale discount perceived by strong-bank investors more prominently. The reason is that strong banks start to experience runs when weak banks are already forced to liquidate a lot of their assets; that is, strong banks are under runs when liquidity conditions are particularly dire. A liquidity injection therefore provides a greater relief to strong-bank investors, thereby boosting heterogeneity and stabilizing the financial system.

Different from ring-fencing and disclosure policies that influence financial fragility only through changing heterogeneity, liquidity injections also have a direct stabilizing effect—that is, a reduction in the fire-sale discount following a liquidity injection benefits banks even with fixed heterogeneity. A calibration exercise based on data of U.S. banks during the 2007–2008 crisis suggests the indirect effect working through changing bank heterogeneity is nontrivial.

The last policy we explore is liquidity requirements. We consider an extension that allows banks to hold both cash and long-term assets. Cash serves as a liquidity buffer against early withdrawals and can be used to acquire assets sold by other banks. Similar to liquidity injection by regulators, required liquidity buffers improve liquidity conditions and directly enhance the stability of the financial system. In addition to that, since strong banks utilize their liquidity buffers under worse liquidity conditions, the same amount of liquidity buffer has a greater stabilizing effect on strong banks than on weak banks. As a result, liquidity requirements can also enhance stability indirectly by boosting bank heterogeneity.

Literature Our model, featuring interconnected fragile banks, builds on two large strands of literature on financial fragility. The first one studies fragility of financial institutions due to strategic complementarities among investors and panic-driven runs

(Diamond and Dybvig, 1983). Settings with strategic complementarities typically feature multiple equilibria if there is no uncertainty about bank fundamentals, with the same policies potentially having different effects in different equilibria. We use the global games technique (Rochet and Vives, 2004; Goldstein and Pauzner, 2005) to pin down a unique equilibrium. Specifically, bank investors in our model are uncertain about the aggregate fundamental and thus cannot perfectly foresee run situations in their own and other banks. The global games approach allows us to characterize the strategic beliefs of bank investors about run situations in their own banks and other banks in the financial system. This approach is particularly appealing for policy analysis as it ties the endogenous likelihood of a crisis to economic fundamentals, which allows us to make precise predictions about costs and benefits of different policy interventions.

The fact that fragilities of individual banks are interrelated through fire sales connects our paper to the second strand of literature that studies contagion through fire-sale spillovers. Early contributions (Cifuentes, Ferrucci, and Shin, 2005; Diamond and Rajan, 2005) emphasize existence of cross-bank spillovers due to a limited liquidity pool. Uhlig (2010) discusses microfoundations of fire-sale discounts faced by banks. A few papers examine the coordination problem across homogeneous financial institutions induced by fire-sale externality. For example, Parlatore (2016) shows that sponsors' support of money market funds complements each other because lower asset prices caused by fire sales make support costlier; Kuong (2021) demonstrates that a tightening of collateral requirements by lenders causes borrowers to default, which results in fire-sale discounts for collateral. Different from these papers, our model features both cross- and within-bank strategic complementarities that intensify each other. More closely related to us are works by Liu (2016, 2018) and Eisenbach (2017) who study models in which bank runs are connected through secondary markets in the absence of uncertainty about the liquidation price. However, their focus is different from ours: Eisenbach (2017) studies the effectiveness of rollover risk as a market disciplinary device in the presence of fire-sale spillovers, and Liu (2016, 2018) shows that fire-sale spillovers might lead to equilibrium multiplicity. We instead emphasize heterogeneous effects of the feedback loop between bank runs and fire sales on different banks in the financial system, and uncover bank heterogeneity as a key determinant of the feedback strength and thus fragility.[6]

A few papers emphasize that regulators should target agents that impose a stronger externality on others in one-complementarity coordination games. Sákovics and Steiner (2012) consider a coordination game with heterogeneous agents and study optimal targeted policies. Shen and Zou (2020) propose policies that screen agents based on their heterogeneous information in global games and illustrate the efficiency of such policies in targeting agents with medium beliefs. Cong, Grenadier, and Hu (2020) argue that saving small banks is cheaper and can generate stronger informational externalities. Probably the closest to us is Choi (2014). He argues that regulators should support strong banks because the fragility of strong banks affects weak banks on the margin but not vice versa. In our setting, fragilities of weak and strong banks always affect each other simultaneously.[7] We show that the reinforcement between within- and cross-bank complementarities, which is absent in Choi (2014), makes heterogeneity beneficial for all banks. This economic mechanism is behind our novel policy analyses. In particular, we show that even broad-based policies such as liquidity injections and liquidity requirements can boost heterogeneity and enhance financial stability.

Our paper is also related to a series of studies on the role of asset commonality for systemic risk (Shaffer, 1994; Acharya, 2009; Stiglitz, 2010; Ibragimov et al., 2011; Wagner, 2010 and 2011; Allen, Babus, and Carletti, 2012; Cabrales et al., 2017; Kopytov, 2023).[8] As discussed above, these papers focus on correlated fundamental bank defaults. They emphasize a trade-off between losses due to failures of individual institutions and systemic crises, which means that whether asset diversification is desirable might depend on the distribution of shocks. For example, Cabrales et al. (2017) argue that fully diversified

[6]Settings with multilayered complementarities have been used to study other questions, e.g. capital and liquidity regulation (Carletti, Goldstein, and Leonello, 2020) and twin crises (Goldstein, 2005). Given their different focus, these papers do not talk about heterogeneity and its importance for fragility.

[7]The difference arises because Choi (2014) uses a binary payoff structure similar to Morris and Shin (1998), while we use Diamond and Dybvig (1983)-like payoffs. In addition, he shows that, depending on the parameters, it can be optimal to support both types of banks or only weak banks. His main result emerges in the continuous-time model where the noise in investors' signals has bounded support.

[8]A recent paper by Song and Thakor (2022) shows how an endogenously arising interbank market can make ex-ante identical banks ex-post different and improve intermediation efficiency.

systems are socially desirable if negative shocks are expected to be small, but asset dispersion is beneficial if shocks are large. In contrast, we focus on panic-driven runs. We show that what matters for financial stability is bank heterogeneity, that is, the dispersion of runs across banks, rather than the dispersion of asset returns. Furthermore, in our framework dispersion in asset returns can reduce run probability for all banks irrespective of distributional assumptions on shocks. This is consistent with empirical findings of Huang et al. (2009) who document that an increase in asset correlation among large U.S. banks leads to an increase in their individual default probabilities.

More generally, our paper makes a theoretical contribution to the global games literature pioneered by Carlsson and van Damme (1993) and later developed into various topics such as currency attacks (Morris and Shin, 1998; Hellwig, Mukherji, and Tsyvinski, 2006) and bank runs (Rochet and Vives, 2004; Goldstein and Pauzner, 2005). A key feature of a standard global games setting is the Laplacian property (Morris and Shin, 2003): A marginal investor is uninformed about the rank of her signal, and her belief about the mass of runners is uniform. Even when investors have heterogeneous payoffs, a version of the Laplacian property holds for the weighted average belief of marginal investors of different types (Sákovics and Steiner, 2012). In those settings, increasing heterogeneity has no impact on investors' run decisions on average, as the optimism of weak-bank investors about the actions of strong-bank investors is exactly offset by the pessimism of strong-bank investors. This is no longer true in our model featuring two types of interacting strategic complementarities. In our setting, the overall fragility depends on the weighted average of beliefs about the interaction between fire-sale pressure and runs in individual banks. These interaction terms are not symmetric across bank types and depend on the degree of heterogeneity.

The remainder of the paper is organized as follows. Section 2 lays out the model. Section 3 presents the main theoretical results on the relationship between heterogeneity and fragility. Section 4 discusses how policies affect heterogeneity and, in turn, stability and welfare. Section 5 concludes.

2 Model

The economy is populated with three types of risk-neutral agents: banks, bank investors, and outside investors. There are three periods, $t = 0, 1, 2$, and no time discounting.

2.1 Banks

There is a continuum of banks indexed by $i \in [0, 1]$. At $t = 0$, bank i collects one unit of capital from a unit mass of investors in the form of demandable debt and makes long-term investment that generates a gross return of $z_i = \theta + \zeta_i$ at $t = 2$, where θ is the aggregate component shared by all banks and ζ_i is the bank-specific component.[9] The cumulative distribution function of the aggregate fundamental $F_\theta(\cdot)$ has a support $[\underline{\theta}, \bar{\theta}]$, where $\infty \geq \bar{\theta} > \underline{\theta} > 0$. The bank-specific shock ζ_i is a zero-mean random variable that takes values $\eta \geq 0$ and $-\eta$ with equal probabilities.[10] The size of bank-specific shocks is restricted to be such that overall productivity is always positive for all banks, i.e., $\underline{\theta} - \eta > 0$. Both θ and ζ_i are unknown at $t = 0$.

Upon shock realizations at $t = 1$ banks become heterogeneous. There are two groups of banks: *strong* banks with $\zeta_i = \eta$ and *weak* banks with $\zeta_i = -\eta$. The masses of the two groups are identical. Parameter η governs the ex-post difference between performances of strong and weak banks and reflects dispersion in bank asset returns, or simply asset dispersion. From an ex-ante perspective, η affects the pairwise correlation between bank fundamentals, i.e., $\text{Corr}(z_i, z_j) = \frac{\mathbb{V}\theta}{\mathbb{V}\theta + \mathbb{V}\zeta_i} = \frac{\mathbb{V}\theta}{\mathbb{V}\theta + \eta^2}$, where $\mathbb{V}$ denotes variance. As η increases, the relative importance of aggregate shocks declines and asset returns of banks i and j become less correlated.

At $t = 1$, after both aggregate and bank-specific productivities are realized, bank investors may choose to withdraw their funds early. Throughout most of the paper, we assume that banks invest all their capital in long-term projects, and so any early with-

[9] Throughout the paper we take the contracts issued by banks to investors as given. Specifically, unlike Diamond and Dybvig (1983), we do not microfound why banks issue demandable debt. The focus of our paper is therefore on within- and cross-bank interactions of bank investors given the contract structure observed in reality.

[10] In Appendix C.5, we show that our main results extend to the case in which ζ_i is a zero-mean random variable that can take an arbitrary number of values with arbitrary probabilities.

drawals from a given bank force this bank to liquidate some of its long-term investment in the asset market. Section 4.4 and Appendix C.7 consider extensions in which banks can hold liquid assets that can be used to repay early withdrawals.

2.2 Outside investors and the asset market

At $t = 1$, if a mass m_i of investors withdraw their funds early from bank i, bank i needs to raise funds of amount m_i by partially liquidating its long-term investment. This means that bank i has to liquidate $\frac{m_i}{p_i}$ units of its long-term investment, where p_i is an endogenous liquidation price.

The asset market is competitive and populated with a unit mass of identical outside investors. Reminiscent of the cash-in-the-market pricing (Allen and Gale, 1994), liquidity is scarce in the asset market, which can cause asset prices to fall below their fundamental values. In particular, in order to purchase a portfolio $\{k_i\}_{i\in[0,1]}$ of bank assets, an outside investor has to raise $L = \int p_i k_i di$ units of external funds by incurring a financing cost. We summarize the financing cost with a simple functional form $g(L) \geq L$, where the spread $g(L) - L$ represents the amount of real losses due to financial imperfections such as agency costs (Gomes, 2001). We assume that $g(\cdot)$ is an increasing and convex function with $g(0) = 0$ and $g'(0) = 1$. The convexity captures an increasing marginal cost of liquidity.[11]

Definition 1. *Given masses of runners* $\boldsymbol{m} = \{m_i\}_{i\in[0,1]}$ *and bank fundamentals* $\boldsymbol{z} = \{z_i\}_{i\in[0,1]}$*, an equilibrium in the asset market consists of outside investors' demand functions* $\{k_i(\boldsymbol{p}, \boldsymbol{z})\}_{i\in[0,1]}$ *and liquidation prices* $\boldsymbol{p} = \{p_i(\boldsymbol{m}, \boldsymbol{z})\}_{i\in[0,1]}$ *such that:*

1. *Given the liquidation prices* $\boldsymbol{p}$*, outside investors' demand functions* $\{k_i(\boldsymbol{p}, \boldsymbol{z})\}_{i\in[0,1]}$ *maximize their expected payoffs:* $\max_{k_i, i\in[0,1]} \int z_i k_i di - g\left(\int p_i k_i di\right)$.

2. *The liquidation prices satisfy the market-clearing conditions:* $m_i = p_i k_i \; \forall i \in [0, 1]$.

The key feature of the asset market is that fire-sale externalities can spill over across banks. One interpretation of this feature is that banks face the same group of buyers

[11] Appendix C.1 provides an alternative microfoundation for fire-sale losses—that is, outside investors are less efficient than banks in managing assets or incur inventory costs when holding them (Shleifer and Vishny, 1992; Kiyotaki and Moore, 1997). We show that that the pricing function in that case takes the same form as in Lemma 1.

of their assets. Even if asset markets for different banks are separated, arbitrage capital might flow across these markets, leading to comoving fire-sale discounts.

The following lemma summarizes the key properties of the liquidation prices.

Lemma 1. *Given masses of runners* $\boldsymbol{m}$ *and bank fundamentals* $\boldsymbol{z}$*, the equilibrium liquidation price for bank i's assets is* $p_i = p(z_i, m) = z_i/\lambda(m)$ $\forall i \in [0,1]$*, where* $m \equiv \int m_i di$ *is the total mass of runners in the economy and* $\lambda(m) \equiv g'(m)$ *is a strictly increasing function.*

Proof. See Appendix B.1. □

The liquidation prices of bank assets are proportional to their productivities z_i's and are subject to a common fire-sale discount factor $\lambda(m)$. The discount factor $\lambda(m)$ increases in the total mass of runners in the entire financial system. Intuitively, if more bank investors withdraw their funds early, banks have to raise more liquidity in the asset market. Because the marginal cost of liquidity is an increasing function, i.e. $g''(\cdot) > 0$, the price discount factor $\lambda(m)$ increases if more bank investors withdraw early. This is akin to the key property of cash-in-the-market pricing: asset prices fall as the liquidity demand exceeds what is available in the market.[12]

2.3 Bank investors and runs

This section describes the behavior of bank investors. For each bank i, there is a unit mass of infinitesimal investors indexed by $l \in [0,1]$. At $t = 0$, each investor contributes one unit of capital to her bank. At that point, bank-specific fundamentals are unknown and so investors are indifferent regarding which bank to invest in.

At $t = 1$, an investor l of bank i observes the bank-specific fundamental ζ_i and receives

[12]In Allen and Gale (1994), $g(L) = \infty$ if L exceeds a certain threshold $\bar{L}$, so outsiders are constrained by the amount of their cash reserves. We allow outsiders to raise additional funds when they run out of cash, so that banks' demand for liquidity can always be satisfied.

a noisy private signal s_{il} about the aggregate fundamental θ,[13]

$$s_{il} = \theta + \sigma\epsilon_{il}. \tag{1}$$

The signal noise ϵ_{il} has a cumulative distribution function $F_\epsilon(\cdot)$, which is differentiable and strictly increasing on its support $[\underline{\epsilon}, \bar{\epsilon}]$. The corresponding probability density function is denoted by $f_\epsilon(\cdot)$. In what follows, we work with a bounded noise support, $-\infty < \underline{\epsilon} < \bar{\epsilon} < \infty$, but our analyses carry through if it is unbounded. The signal noise ϵ_{il} has a zero mean and a unit variance, i.e. $\mathbb{E}\epsilon_{il} = 0$ and $\mathbb{V}\epsilon_{il} = 1$, so that σ is the standard deviation of the private signal conditional on the aggregate fundamental θ. This information structure follows a conventional global games setup.

With probability $\bar{m} \in (0, 1)$, investor l is "non-sleepy" and may withdraw her funds from her bank at $t = 1$.[14] With probability $1 - \bar{m}$, investor l is "sleepy" and neglects the option to withdraw early. Therefore, bank i needs to liquidate at most a fraction $\frac{\bar{m}}{p_i}$ of its assets if all "non-sleepy" investors withdraw their funds early. We rule out bank failures by assuming that $\frac{\bar{m}}{p_i} \leq 1$ (as in Chen, Goldstein, and Jiang, 2010). Goldstein and Pauzner (2005) show that the possibility of bank failure creates a region of strategic substitution, making the analysis much more technically involved. Importantly, while the assumption on investors' limited attention rules out complete bank failures, the key property of runs is preserved—they reduce the amount of funds to stayers and are socially undesirable because they trigger fire sale losses. In Appendix C.4, we show that our main results hold even if bank failures are possible.

In the absence of bank failures, investors who withdraw early are guaranteed to get their funds back at $t = 1$. At $t = 2$, the investment return to bank i is equally distributed among investors who have not withdrawn their funds at $t = 1$. A "non-sleepy" investor

[13] Section 4.2 extends the model by allowing for partially informative signals about bank-specific fundamentals. Under imperfect information about the aggregate fundamental θ, there exist strategic uncertainties both within and across banks, which allows investors to coordinate and supports equilibrium uniqueness. Under perfect information about θ, multiplicity is possible (Liu, 2016).

[14] We assume that investors who withdraw early do not redeposit their funds to other banks. Appendix C.6 considers a simple extension in which redepositing is allowed. It shows, in particular, that our main results hold if *some* investors who withdraw early do not redeposit their funds to other banks.

l's payoff conditional on her decision to withdraw early $a_{il} \in \{run,\ stay\}$ from bank i is

$$u_i(a_{il}) = \begin{cases} 1 & \text{if } a_{il} = run, \\ \dfrac{z_i\left(1 - \frac{m_i}{p_i}\right)}{1 - m_i} & \text{if } a_{il} = stay. \end{cases}$$

Plugging in the market-clearing liquidation price p_i derived in Lemma 1, we can express the incremental payoff from staying as

$$\pi(z_i, m_i, m) \equiv u_i\,(stay) - u_i\,(run) = \frac{z_i - m_i\lambda(m)}{1 - m_i} - 1. \tag{2}$$

In particular, a "non-sleepy" investor l of bank i runs if and only if the expected incremental payoff given her signal is negative,[15]

$$\mathbb{E}[\pi(z_i, m_i, m)|s_{il}] < 0. \tag{3}$$

Equation (2) reveals two types of strategic complementarities featured by our model. First, there is a *within-bank strategic complementarity*: an investor's incremental payoff from staying declines if more investors in her own bank run.[16] On top of that, the fire-sale externalities in the asset market give rise to a *cross-bank strategic complementarity*: an investor's incremental payoff from staying declines if more investors in other banks run. More importantly, these two complementarities amplify each other, namely,

$$\frac{\partial^2 \pi(z_i, m_i, m)}{\partial m_i \partial m} = -\frac{\lambda'(m)}{(1 - m_i)^2} < 0. \tag{4}$$

Banks that encounter more runs (higher m_i's) have to liquidate more assets, which naturally makes the payoff to staying investors more sensitive to a reduction in the liquidation price and therefore the total mass of runners in the economy m. In Section 3.3, we discuss this feature in detail and explain its importance for our results.

[15] We assume that bank investors make their withdrawal decisions before liquidation prices are realized. As a result, bank investors only receive exogenous private signals (1) about the aggregate fundamental. Consequently, although liquidation prices are endogenous in our model, they do not serve as endogenous public signals to bank investors (as in e.g. Angeletos and Werning, 2006).

[16] Note that when $\lambda(m) < z_i$, the fire-sale discount is small relative to long-term asset return, so the incremental payoff function (2) implies within-bank strategic substitution. However, as we show in Appendix D, the incremental payoff function satisfies a single-crossing property, which, together with mild assumptions on the noise distribution, are sufficient for the uniqueness of threshold equilibrium.

2.4 Timeline and equilibrium definition

Figure 1 depicts the timeline of our model.

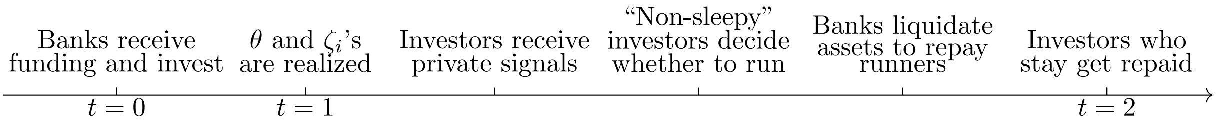

Figure 1: Timeline

Denote by $\{a_{il}(s_{il}, \zeta_i)\}_{i,l\in[0,1]}$ the set of strategies of "non-sleepy" investors that map bank-specific fundamentals ζ_i and their private signals about the aggregate fundamental s_{il} to their action space $a_{il} \in \{run, stay\}$.

Definition 2. *Bank investors' strategies* $\{a_{il}(s_{il}, \zeta_i)\}_{i,l\in[0,1]}$*, outside investors' demand functions* $\{k_i(\boldsymbol{p}, \boldsymbol{z})\}_{i\in[0,1]}$*, liquidation prices* $\boldsymbol{p} = \{p_i(\boldsymbol{m}, \boldsymbol{z})\}_{i\in[0,1]}$*, and masses of runners* $\boldsymbol{m} = \{m_i(\theta, \zeta_i)\}_{i\in[0,1]}$ *constitute an equilibrium if*

1. *Given* $\boldsymbol{m}$ *and* $\boldsymbol{z}$*,* $\{k_i(\boldsymbol{p}, \boldsymbol{z})\}_{i\in[0,1]}$ *and* $\boldsymbol{p}$ *constitute an equilibrium in the asset market as in Definition 1;*

2. *Given her private signal and the bank-specific fundamental of her bank, each "non-sleepy" investor forms beliefs about* $\boldsymbol{p}$ *and* $\boldsymbol{m}$ *and runs if and only if* (3) *holds;*

3. $m_i(\theta, \zeta_i) = \int \mathbb{1}\{a_{il}(s_{il}, \zeta_i) = run\}\, dl$.

2.5 Global games and threshold equilibrium

We focus on threshold equilibria in which investor l of bank i runs when $s_{il} < \theta_i^*$ and stays when $s_{il} \geq \theta_i^*$.[17] We characterize the threshold equilibrium in this section and prove its uniqueness in Appendix D. Throughout most of the paper, we focus on the limiting case of infinitely precise signals, $\sigma \to 0$, as is typical in the global games literature. As such, run thresholds should be understood as limits with $\sigma \to 0$. However, for brevity and when it does not cause confusion, we do not explicitly write endogenous variables

[17]Because the incremental payoff function (2) does not always imply within-bank strategic complementarity, non-threshold equilibria cannot be ruled out. However, under additional assumptions on the noise distribution—e.g., if it is uniform—it can be shown that non-threshold equilibria do not exist.

as limits. For example, in most cases we write a bank i's run threshold as θ_i^*, not as $\lim_{\sigma\to 0}\theta_i^*(\sigma)$. We show that our results are robust to non-negligible noise in Appendix C.3.

A threshold investor with a signal $s_{il} = \theta_i^*$ is indifferent between running and staying,

$$\int_0^1 \frac{\theta(x) + \zeta_i - \lambda\left(m_i^{tot}(x)\right)\bar{m}x}{1-\bar{m}x}dx = 1, \tag{5}$$

where x is the fraction of "non-sleepy" investors running on bank i. As is standard in global games models, a threshold investor with a signal θ_i^* has a Laplacian belief. That is, she believes that the mass of runners within her own bank, $\bar{m}x = \bar{m}F_\epsilon\left(\frac{\theta_i^*-\theta}{\sigma}\right)$, is uniformly distributed (Morris and Shin, 2003). Crucially, because her signal is informative about the aggregate productivity, she also makes inference about actions of investors of other banks. Given her belief about the fraction of "non-sleepy" investors running on her bank x, her belief about the aggregate fundamental is $\theta(x) = \theta_i^* - \sigma F_\epsilon^{-1}(x)$ and the amount of runs on bank j is $\bar{m}F_\epsilon\left(\frac{\theta_j^*-\theta(x)}{\sigma}\right)$. The total mass of runners in the economy from the perspective of this investor is then

$$m_i^{tot}(x) = \bar{m}\int F_\epsilon\left(\frac{\theta_j^*-\theta(x)}{\sigma}\right)dj = \bar{m}\int F_\epsilon\left(\frac{\theta_j^*-\theta_i^*}{\sigma} + F_\epsilon^{-1}(x)\right)dj. \tag{6}$$

Taking the limit $\sigma \to 0$ in (5) and using (6), we can write thresholds θ_i^* for investors of all banks in the economy. Given the binary structure of bank-specific shocks, the run thresholds of strong and weak banks, denoted as θ_s^* and θ_w^*, are determined by

$$\theta_s^* + \eta = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + I_s(\Delta)\right), \tag{7}$$

$$\theta_w^* - \eta = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + I_w(\Delta)\right), \tag{8}$$

where

$$I_s(\Delta) \equiv \int_0^1 \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right)\right)\frac{\bar{m}x}{1-\bar{m}x}dx, \tag{9}$$

$$I_w(\Delta) \equiv \int_0^1 \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\right)\frac{\bar{m}x}{1-\bar{m}x}dx, \tag{10}$$

and

$$\Delta \equiv \lim_{\sigma \to 0} \frac{\theta_w^* - \theta_s^*}{\sigma}. \tag{11}$$

Here $I_s(\Delta)$ and $I_w(\Delta)$ represent expected fire-sale losses borne by threshold investors with Laplacian beliefs in strong and weak banks, respectively. In what follows, we refer to $I_s(\Delta)$ and $I_w(\Delta)$ as fire-sale pressure on strong and weak banks. Notably, fire-sale pressure on a given bank is exerted by all banks, including banks of its own group.

The endogenous variable Δ is the distance between the two run thresholds. It measures the extent to which runs (and, hence, tail risks) are dispersed across strong and weak banks. Therefore, we call Δ *bank heterogeneity*, or simply heterogeneity. It is straightforward to verify that $\Delta \geq 0$.[18] If $\Delta = 0$, the financial system is homogeneous—that is, each threshold investor believes that all banks in the economy face the same amount of runs. If $\Delta > 0$, runs are heterogeneous, and threshold investors of strong (weak) banks believe that weak (strong) banks face more (less) severe run problems.

Bank heterogeneity Δ governs cross-bank interactions in this economy. Because banks impose fire-sale externalities on one another, investors need to evaluate the run situations in all banks simultaneously to make their own decisions whether to withdraw early. Consider a threshold investor of a strong bank as an example. As shown in Equation (9), if a fraction x of "non-sleepy" investors run on strong banks, she expects a fraction $F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) \geq x$ of "non-sleepy" investors to run on weak banks. Δ plays an important role in shaping investors' beliefs about the total amount of runs in the economy and thus the fire-sale discount, which in turn governs their own run decisions.

3 Heterogeneity and stability

The key premise of our model is that within-bank run problems and cross-bank complementarities due to fire-sale spillovers are mutually reinforcing. This observation is by no means novel in the literature.[19] The main novelty of our paper is that the strength

[18] To see that, suppose that $\Delta < 0$ and hence $\theta_w^* < \theta_s^*$. Then Equations (9) and (10) imply that $I_w(\Delta) \geq I_s(\Delta)$. From Equations (7) and (8) and the fact that $\eta \geq 0$ it then follows that $\theta_w^* \geq \theta_s^*$, which is a contradiction.

[19] See, among others, Parlatore (2016), Liu (2016, 2018), Eisenbach (2017), Kuong (2021).

of the complementarity reinforcement—and hence the overall financial fragility—depends on the structure of the financial system. Our main insight is that differentiating banks moderates the adverse effects of the complementarity reinforcement and thus stabilizes individual banks and the financial system as a whole. To emphasize the importance of heterogeneity, we start by considering a homogeneous benchmark in Section 3.1 and then analyze the role of heterogeneity in Sections 3.2 and 3.3.

3.1 Homogeneous benchmark

Consider the homogeneous benchmark in which all banks have the same bank-specific productivities, that is, $\eta = 0$. Bank heterogeneity is then $\Delta = 0$, and investors of all banks follow the same threshold strategy with the threshold θ_0^*. From (5), this threshold is implicitly given by

$$\int_0^1 \frac{\theta_0^* - \lambda(\bar{m}x)\bar{m}x}{1-\bar{m}x} dx = 1. \tag{12}$$

Even in the absence of heterogeneity, complementarity reinforcement is at play and runs on individual banks are aggravated by the fire-sale externalities. To illustrate this point, we consider an alternative setting in which the fire-sale discount is fixed at the average level $\bar{\lambda} = \int_0^1 \lambda(\bar{m}x)dx$, so that there are no complementarity interactions. In this case, the run threshold $\hat{\theta}_0^*$ is implicitly given by

$$\int_0^1 \frac{\hat{\theta}_0^* - \bar{\lambda}\bar{m}x}{1-\bar{m}x} dx = 1. \tag{13}$$

Comparing Equations (12) and (13), we obtain the following.

$$\begin{aligned}\left(\theta_0^* - \hat{\theta}_0^*\right) \int_0^1 \frac{1}{1-\bar{m}x} dx = \int_0^1 \frac{\lambda(\bar{m}x)\bar{m}x}{1-\bar{m}x} dx - \int_0^1 \frac{\bar{\lambda}\bar{m}x}{1-\bar{m}x} dx = \\ \mathbb{E}\left[\lambda(\bar{m}x)\frac{\bar{m}x}{1-\bar{m}x}\right] - \mathbb{E}\left[\lambda(\bar{m}x)\right]\mathbb{E}\left[\frac{\bar{m}x}{1-\bar{m}x}\right] = \\ \mathrm{Cov}\left(\lambda(\bar{m}x), \frac{\bar{m}x}{1-\bar{m}x}\right) > 0,\end{aligned}$$

where expectations and covariance are taken with respect to a random variable x that is uniformly distributed on $[0, 1]$, and where the covariance is positive because $\lambda(\bar{m}x)$ and $\frac{\bar{m}x}{1-\bar{m}x}$ are both increasing functions of x.

We can see that the model with the complementarity interaction features a higher fragility, $\theta_0^* > \hat{\theta}_0^*$. With the complementarity reinforcement, runs are particularly detrimental to stayers when the fire-sale discount is high. If a threshold investor expects the within-bank run problem to be severe and the fire-sale discount to be large in the same states of the world, her expected payoff from staying is low. If, in contrast, within-bank runs and the fire-sale discount are uncorrelated—as is the case in the alternative model without complementarity interaction—a threshold investor is more likely to stay.

This simple example illustrates that the interacting within- and cross-bank complementarities are detrimental for stability in a homogeneous financial system. In what follows, we show that heterogeneity can be helpful to alleviate the complementarity reinforcement. In particular, we emphasize the importance of bank heterogeneity Δ and analyze how various model primitives affect Δ and, ultimately, financial stability.

3.2 Heterogeneity and stability: Characterizing the relation

We decompose our analyses into two cases. In the first case, bank heterogeneity Δ is not too large, $\Delta < \bar{\Delta} \equiv \bar{\epsilon} - \underline{\epsilon}$, and the financial system features nontrivial cross-bank strategic uncertainties. In this case, threshold investors of strong banks are uncertain about run behavior of weak-bank investors, and vice versa. Formally speaking, it implies that the fire-sale pressure terms $I_s(\Delta)$ and $I_w(\Delta)$, given by (9) and (10), depend on Δ. In the second case, $\Delta \geq \bar{\Delta}$, and the fire-sale pressure terms are constants. In this case, threshold strong-bank investors are certain that *all* "non-sleepy" weak-bank investors are running, and threshold weak-bank investors are certain that *none* of strong-bank investors are running. We think of such a lack of cross-bank strategic uncertainty to be less empirically relevant, as in practice it is hard for investors to precisely assess the run situations in other banks during financial turmoils. In comparison, the first case is more empirically relevant and also more interesting from a theoretical standpoint.

Nontrivial cross-bank strategic interactions Using Equations (7) and (8), we can write average fragility $\theta^* \equiv \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$ as

$$\theta^* = \frac{1}{\int_0^1 \frac{dx}{1-\bar{m}x}} \left(1 + \frac{1}{2}I_s(\Delta) + \frac{1}{2}I_w(\Delta)\right), \tag{14}$$

where bank heterogeneity Δ is implicitly given by

$$\eta = \frac{1}{2\int_0^1 \frac{dx}{1-\bar{m}x}} \left(I_s(\Delta) - I_w(\Delta)\right). \tag{15}$$

By definition of bank heterogeneity in the limit of negligible information friction (11), a finite Δ implies that the run thresholds of weak and strong banks are infinitely close to each other, i.e. $\lim\limits_{\sigma\to 0}\theta^* = \lim\limits_{\sigma\to 0}\theta_s^* = \lim\limits_{\sigma\to 0}\theta_w^*$. The result that individual run thresholds are infinitely close is a typical feature of global games with heterogeneous players and infinitely precise signals (Frankel, Morris, and Pauzner, 2003).[20]

The fact that θ_s^* and θ_w^* are infinitely close to each other implies that, from the $t = 0$ point of view, fragilities of both weak and strong banks are characterized by just one threshold θ^*. Therefore, θ^* measures the ex-ante fragility of all banks in this economy. Importantly, although θ_s^* and θ_w^* are infinitely close to each other, there can still be a nontrivial difference between them from investors' perspective at the moment of runs. This is because investors receive infinitely precise signals about the aggregate fundamental at $t = 1$ and thus can distinguish between θ_s^* and θ_w^* even if they are infinitely close. As a result, bank heterogeneity Δ, which by (11) is the distance between θ_w^* and θ_s^* in units of σ, can still be finite. Crucially, as Δ determines the strength of cross-bank strategic interactions, it has a first-order effect on the common run threshold θ^*.

Equation (14) shows how θ^* is affected by bank heterogeneity Δ and model primitives, whereas (15) implicitly defines Δ as a function of model primitives. Naturally, such a formulation allows us to decompose the effect of any regulatory policy on financial stability into two parts. Holding Δ fixed, a change in any policy-related parameter v can have a direct effect on θ^*. This channel captures policy impact holding cross-bank

[20] With non-negligible information friction ($\sigma > 0$), run thresholds of weak and strong banks are distinct and centered around the average threshold, $\theta_s^*(\sigma) < \theta^*(\sigma) < \theta_w^*(\sigma)$. We show in Appendix C.3 how our results extend to this case.

interactions fixed. Furthermore, a policy can affect θ^* indirectly through reshaping cross-bank interactions and thus bank heterogeneity Δ. That is,

$$\frac{d\theta^*}{dv} = \underbrace{\frac{\partial\theta^*}{\partial v}}_{\text{Direct effect}} + \underbrace{\frac{\partial\theta^*}{\partial\Delta}\frac{d\Delta}{dv}}_{\text{Indirect effect}}. \tag{16}$$

The focus of our paper is on the indirect effect, which has received much less attention in the existing literature. The sign of the indirect effect depends on two components. First, it depends on the relationship between heterogeneity and fragility $\frac{\partial\theta^*}{\partial\Delta}$. Proposition 1 below uses Equation (14) to establish that financial systems with higher bank heterogeneity are more stable, namely, $\frac{\partial\theta^*}{\partial\Delta} < 0$. Section 3.3 discusses the key economic force behind this result, that is, reinforcing within- and cross-bank complementarities. Second, the indirect effect depends on how the policy variable affects heterogeneity, that is, on $\frac{d\Delta}{dv}$. Equation (15) reveals various determinants of heterogeneity, such as asset dispersion η and liquidity conditions in the asset market that shape the fire-sale discount faced by strong and weak banks. In Section 4, we examine several regulatory policies and discuss how they affect stability both directly and indirectly.

Proposition 1. *When $\Delta < \bar{\Delta} = \bar{\epsilon} - \underline{\epsilon}$, run thresholds of strong and weak banks are infinitely close to the average threshold θ^*, which captures the fragility of the financial system. Moreover, $\theta^* = \theta^*(\Delta)$ decreases in bank heterogeneity Δ.*

Proof. See Appendix B.2. □

Equation (14) reveals that the common run threshold θ^* is determined by the average fire-sale pressure on strong and weak banks, $\frac{1}{2}I_s(\Delta) + \frac{1}{2}I_w(\Delta)$. Recall that, by definition of heterogeneity (11), Δ is the normalized distance between run thresholds of weak and strong banks. From the perspective of strong-bank investors, weak banks become more fragile as Δ goes up. As a result, strong-bank investors expect a larger fire-sale externality imposed by weak banks. Mathematically, it is captured by the fact that the fire-sale pressure on strong banks $I_s(\Delta)$ increases in Δ. At the same time, in the view of weak-bank investors, strong banks become more stable, and so these investors expect lower fire-sale pressure $I_w(\Delta)$. Proposition 1 shows that the latter effect is more substantial than the former, that is, the average fire-sale pressure and the run threshold θ^* decrease with Δ. Intuitively, this is because weak banks suffer from more severe runs in any state of the world and are therefore more sensitive to changes in the fire-sale pressure than strong banks. Section 3.3 illustrates in detail that the key force behind this result is mutually reinforcing within- and cross-bank complementarities.

It is worth emphasizing that both weak and strong banks become more stable as the difference between their fragilities Δ enlarges. This can be clearly seen in the case with negligible information friction as weak and strong banks have infinitely close run thresholds. In Appendix C.3, we analyze the non-limiting case where the run thresholds of weak and strong banks decouple. In that case, the average fragility still declines in Δ, and both weak and strong banks tend to benefit from a large Δ unless the signal noise is too large.

Trivial cross-bank strategic interactions When the financial system features large heterogeneity, $\Delta \geq \bar{\Delta}$, there are no strategic uncertainties between investors of weak and strong banks. In particular, threshold investors of strong banks are certain that all "non-sleepy" investors of weak banks will run; conversely, threshold investors of weak banks are certain that no investor in strong banks will run. Therefore, when investors make their run decisions, they only need to make inferences about the behavior of investors in the same type of banks. In this case, fire-sale pressure on both strong and weak banks, $I_s(\Delta)$ and $I_w(\Delta)$, no longer depend on Δ, and runs in strong and weak banks can be analyzed separately.

Formally, Equations (7) and (8) for the run thresholds become

$$\theta_s^* = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + \int_0^1 \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}\right)\frac{\bar{m}x}{1-\bar{m}x}dx\right) - \eta, \tag{17}$$

$$\theta_w^* = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + \int_0^1 \lambda\left(\frac{\bar{m}}{2}x\right)\frac{\bar{m}x}{1-\bar{m}x}dx\right) + \eta. \tag{18}$$

Because the run thresholds of weak and strong banks do not vary with Δ, the average fragility of banks θ^* does not vary with Δ either.[21]

[21] In the case of trivial cross-bank strategic interactions, θ_s^* and θ_w^* are not infinitely close. Then by definition (11), $\Delta \to \infty$ in the limit of infinitely small σ. If σ is not infinitely small, then Δ is finite even in the case of trivial strategic interactions.

3.3 Role of reinforcing complementarities

The key force behind Proposition 1 is that the within- and cross-bank complementarities are mutually reinforcing, which in the model is captured by (4). Weak banks are more sensitive to changes in the fire-sale discount because they experience more runs and need to liquidate more assets. As we discussed in the previous section, when $\Delta < \bar{\Delta}$, an increase in Δ alleviates cross-bank fire-sale pressure on weak banks and worsens that on strong banks. However, since weak banks are more sensitive to the change, the benefit for weak banks outweighs the loss for strong banks, and the overall stability increases. In this section, we illustrate this point formally by comparing marginal impacts of Δ on $I_w(\Delta)$ and $I_s(\Delta)$.

First, consider the fire-sale pressure on strong banks. Equation (9) can be written as

$$I_s(\Delta) = \int_{\underline{\epsilon}}^{\bar{\epsilon}-\Delta} \lambda\left(\frac{\bar{m}}{2}F_\epsilon(\epsilon_s) + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + \epsilon_s\right)\right) \frac{\bar{m}F_\epsilon(\epsilon_s)}{1-\bar{m}F_\epsilon(\epsilon_s)} dF_\epsilon(\epsilon_s) + \int_{\bar{\epsilon}-\Delta}^{\bar{\epsilon}} \lambda\left(\frac{\bar{m}}{2}F_\epsilon(\epsilon_s) + \frac{\bar{m}}{2}\right) \frac{\bar{m}F_\epsilon(\epsilon_s)}{1-\bar{m}F_\epsilon(\epsilon_s)} dF_\epsilon(\epsilon_s),$$

where we change the variable of integration $x = F_\epsilon(\epsilon_s)$. Here, ϵ_s represents the realization of signal noise for a threshold investor of a strong bank. From (1), $\epsilon_s = \frac{\theta_s^* - \theta}{\sigma}$ for the strong-bank investor receiving a threshold signal θ_s^*. Because investors receiving worse signals withdraw their funds prematurely, the mass of runners in each strong bank is $\bar{m}F_\epsilon(\epsilon_s)$. Similarly, the mass of runners on each weak bank is $\bar{m}F_\epsilon(\Delta + \epsilon_s)$.

An increase in heterogeneity bolsters the fire-sale pressure on strong banks,

$$\frac{\partial I_s}{\partial \Delta} = \int_{\underline{\epsilon}}^{\bar{\epsilon}-\Delta} \underbrace{f_\epsilon(\epsilon_s)}_{A_{s,1}} \times \underbrace{\frac{\bar{m}}{2} f_\epsilon(\Delta+\epsilon_s)}_{A_{s,2}} \times \underbrace{\lambda'\left(\frac{\bar{m}}{2}F_\epsilon(\epsilon_s) + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + \epsilon_s\right)\right) \frac{\bar{m}F_\epsilon(\epsilon_s)}{1-\bar{m}F_\epsilon(\epsilon_s)}}_{A_{s,3}} d\epsilon_s > 0. \quad (19)$$

The integrand can be decomposed into three parts. First, $A_{s,1}$ represents the probability of a state in which a threshold investor receives a signal with a noise realization ϵ_s. As Δ goes up, this investor expects a larger mass of runners on weak banks, which is captured by the second term $A_{s,2}$. More runs on weak banks increase the fire-sale discount and reduce the payoff to the threshold investor. The third term $A_{s,3}$ captures the marginal decrease in the payoff. Integrating over all possible states yields the total change in the fire-sale pressure on a strong bank in response to a marginal increase in Δ.

At the same time, an increase in heterogeneity alleviates the fire-sale pressure on weak banks,

$$\frac{\partial I_w}{\partial \Delta} = -\int_{\underline{\epsilon}+\Delta}^{\bar{\epsilon}} f_\epsilon(\epsilon_w)\frac{\bar{m}}{2} f_\epsilon(-\Delta+\epsilon_w)\lambda'\left(\frac{\bar{m}}{2}F_\epsilon\left(-\Delta+\epsilon_w\right)+\frac{\bar{m}}{2}F_\epsilon(\epsilon_w)\right)\frac{\bar{m}F_\epsilon(\epsilon_w)}{1-\bar{m}F_\epsilon(\epsilon_w)}d\epsilon_w =$$
$$-\int_{\underline{\epsilon}}^{\bar{\epsilon}-\Delta} \underbrace{f_\epsilon(\Delta+\epsilon_s)}_{A_{w,1}} \times \underbrace{\frac{\bar{m}}{2} f_\epsilon(\epsilon_s)}_{A_{w,2}} \times \underbrace{\lambda'\left(\frac{\bar{m}}{2}F_\epsilon(\epsilon_s)+\frac{\bar{m}}{2}F_\epsilon\left(\Delta+\epsilon_s\right)\right)\frac{\bar{m}F_\epsilon(\Delta+\epsilon_s)}{1-\bar{m}F_\epsilon(\Delta+\epsilon_s)}}_{A_{w,3}} d\epsilon_s < 0, \quad (20)$$

where the second equality is obtained by changing the variable of integration, $\epsilon_s = \epsilon_w - \Delta$. Same as (19) for the strong banks, the integrand can be decomposed into three parts.

Compare the magnitudes of a marginal impact of Δ on strong and weak banks, that is, the absolute values of (19) and (20). For a given realization of the aggregate fundamental θ, if a threshold investor of a strong bank has a signal noise realization $\epsilon_s = \frac{\theta_s^*-\theta}{\sigma}$, a threshold investor of a weak bank must have a noise realization $\epsilon_w = \frac{\theta_w^*-\theta}{\sigma} = \Delta + \epsilon_s$. These two investors hold the same beliefs about the run situations in the economy: The mass of runs in all strong and all weak banks are $\frac{\bar{m}}{2}F_\epsilon(\epsilon_s)$ and $\frac{\bar{m}}{2}F_\epsilon(\epsilon_s+\Delta)$, respectively. From the perspective of these investors, changes in the masses of runners in response to an increase in Δ are symmetric, such that $A_{s,1} \times A_{s,2} = A_{w,1} \times A_{w,2}$. Moreover, since they hold the same view on the aggregate runs in the economy, they expect the same marginal impact on the fire-sale discount, captured by $\lambda'\left(\frac{\bar{m}}{2}F_\epsilon(\epsilon_s)+\frac{\bar{m}}{2}F_\epsilon\left(\Delta+\epsilon_s\right)\right)$. However, the resulting changes in their payoffs are different. Specifically, weak banks experience more runs than strong banks, i.e. $F_\epsilon(\epsilon_w) = F_\epsilon(\Delta+\epsilon_s) > F_\epsilon(\epsilon_s)$. Due to reinforcing within- and cross-bank complementarities, the same change in the fire-sale cost has a more profound effect on weak-bank investors than on strong-bank investors, which implies that $A_{w,3} > A_{s,3}$. Because this is true for any realization of the aggregate fundamental, we have

$$\left|\frac{\partial I_w}{\partial \Delta}\right| > \left|\frac{\partial I_s}{\partial \Delta}\right|.$$

From Equation (14), it then follows that $\frac{\partial \theta^*}{\partial \Delta} < 0$, that is, financial systems in which runs are more dispersed across banks are more stable. This result has important policy implications, as reflected by the indirect effect in Equation (16). The fact that the two types of complementarities are mutually reinforcing plays a decisive role for this result.

In our model, mutually reinforcing complementarities is a natural implication of the existence of asset fire sales and a standard Diamond and Dybvig (1983) payoff structure. The following proposition formally establishes, with a more general payoff function, the importance of mutually reinforcing complementarities for our result in Proposition 1.

Proposition 2. *Consider an investor of bank i whose incremental payoff from staying is $\pi(z_i, m_i, m) = z_i \pi_1(m_i) + \pi_2(m_i, m)$, where $\pi_1(m_i)$ is positive. In any threshold equilibrium, if $\Delta < \bar{\Delta}$, run thresholds of strong and weak banks are infinitely close to the average threshold θ^*. Moreover, if $\frac{\partial^2 \pi}{\partial m \partial m_i} \lesseqqgtr 0$ then $\frac{\partial \theta^*(\Delta)}{\partial \Delta} \lesseqqgtr 0$.*

Proof. See Appendix B.2. □

In the baseline setting, $\pi_1(m_i) = \frac{1}{1-m_i}$ and $\pi_2(m_i, m) = -\frac{\lambda(m) m_i}{1-m_i}$, so we have two complementarities reinforcing each other, $\frac{\partial^2 \pi}{\partial m \partial m_i} < 0$. Appendix C.2 further generalizes Proposition 2 by imposing even milder restrictions on the incremental payoff function.

Note that if $\frac{\partial^2 \pi}{\partial m \partial m_i} = 0$, then Δ does not affect the common threshold θ^*. This result echoes Sákovics and Steiner (2012). They show that in global games with heterogeneous agents, the weighted average belief about the aggregate action—economy-wide amount of runs in our model—is uniformly distributed. Moreover, in the absence of the interaction between the within- and cross-bank complementarities, only this weighted average belief matters for the common run threshold. Heterogeneity therefore does not affect the run threshold.

When the two complementarities do interact, the run threshold depends on the interaction terms between the amounts of runs in the whole economy and within a particular bank. Therefore, the powerful result of Sákovics and Steiner (2012) does not hold in our setting, making the analyses much more cumbersome. By comparing (19) and (20), we can see that the interaction terms are not symmetric across weak and strong banks: a reduction in the fire-sale discount benefits weak banks more than an increase in the fire-sale discount of the same size hurts strong banks. A resulting sizable reduction in fragility of weak banks has a positive effect on strong banks, thus pushing the whole financial system to an equilibrium with a higher financial stability.

4 Regulatory policies and heterogeneity

The previous section shows that heterogeneity has important effects on financial stability in the presence of reinforcing complementarities. Notably, moving from stability to welfare is straightforward in our model. Since all agents are risk-neutral, welfare at $t = 0$ can be measured as the expected output from banks' long-term projects net of the losses due to inefficient fire sales. Specifically,

$$W = \underbrace{\int \theta dF_\theta(\theta)}_{\text{Expected output}} - \underbrace{F_\theta(\theta_s^*)\,(g(\bar{m}) - \bar{m})}_{\text{Losses if all banks are under run}} - \underbrace{\left(F_\theta\left(\theta_w^*\right) - F_\theta\left(\theta_s^*\right)\right)\left(g\left(\frac{1}{2}\bar{m}\right) - \frac{1}{2}\bar{m}\right)}_{\text{Losses if only weak banks are under run}}, \quad (21)$$

where $F_\theta(\cdot)$ is the cumulative distribution function of the aggregate fundamental θ at $t = 0$. If strategic interactions between investors of weak and strong banks are nontrivial and the information friction is negligible, θ_s^* and θ_w^* are infinitely close (Proposition 1), and (21) simplifies to

$$W = \int \theta dF_\theta(\theta) - F_\theta(\theta^*)\,(g(\bar{m}) - \bar{m})\,. \quad (22)$$

Equations (21) and (22) imply that reducing bank fragilities—that is, reducing run thresholds—is welfare-improving. In this section, we look into policies that are widely adopted by regulators in order to make the financial system more resilient. We emphasize that many policies, by affecting heterogeneity, have an indirect effect on stability (see Equation (16)) that has been previously overlooked. Interestingly, as we will show in this section, even though some policies are generally considered broad-based, they can impact certain banks more than others, thus changing heterogeneity.

Section 4.1 studies ring-fencing that affects bank asset dispersion. Section 4.2 extends the model by allowing for imperfect information about bank-specific fundamentals and studies disclosure of bank-specific fundamentals. Section 4.3 examines liquidity injections into the asset market. Section 4.4 allows banks to hold cash to acquire liquidated assets of their peers and examines effects of required liquidity buffers.

4.1 Ring-fencing

Ring-fencing refers to separating large banks' balance sheets and restricting fund reallocations across ring-fenced subsidiaries. It is typically conducted along two dimensions. First, separations can take place according to the service divisions. In the United States, the Volcker Rule restricts proprietary trading by commercial banks, essentially spinning off their investment banking activities. Similarly, starting from 2019, largest U.K. banks are required to separate core businesses in retail banking from investment banking.[22] Second, separations can be carried out according to geographic locations. For example, the Fed requires foreign banking organizations with more than \$50 billion in U.S. subsidiary assets to put all their U.S. subsidiaries under an intermediate holding company (Kreicher and McCauley, 2018). Geographic ring-fencing has also been pursued by the European regulator through imposing heavier restrictions on foreign-owned subsidiaries and restricting intragroup cross-border asset transfers (Enria, 2018).

In the context of our model, a regulator separating banks into subsidiaries with different business or geographic focuses effectively increases the dispersion of asset returns η. For instance, consider splitting a large bank issuing mortgages in New York and San Francisco into two separate banks, one of which operating only in New York while the other one only in San Francisco. As long as regional shocks of New York and San Francisco are not perfectly correlated, the two separate banks are exposed to larger bank-specific shocks than the conglomerate.

The comparative statics of the run thresholds with respect to η are established in Proposition 3 and illustrated in Figure 2. For the rest of this section, we express run thresholds and bank heterogeneity as functions of η.

[22] Ring-fencing was first introduced through the Financial Services (Banking Reform) Act 2013, followed by adjustments in further legislation. See a summary at https://www.gov.uk/government/publications/ring-fencing-information/ring-fencing-information.

Proposition 3. *Define*

$$\bar{\eta} \equiv \frac{1}{2\int_0^1 \frac{dx}{1-\bar{m}x}} \int_0^1 \left[\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}\right) - \lambda\left(\frac{\bar{m}}{2}x\right)\right] \frac{\bar{m}x}{1-\bar{m}x} dx > 0. \tag{23}$$

If $\eta \in (0, \bar{\eta})$, then cross-bank strategic interactions are nontrivial, $\Delta(\eta) < \bar{\Delta}$, and run thresholds of strong and weak banks are infinitely close to $\theta^*(\eta)$. Moreover, $\Delta(\eta)$ increases in η and $\theta^*(\eta)$ decreases in η.

Proof. See Appendix B.3. □

When asset performances of weak and strong banks are not too dispersed, $\eta < \bar{\eta}$, cross-bank interactions are nontrivial, $\Delta < \bar{\Delta}$, and the run thresholds of strong and weak banks are infinitely close to each other. It is straightforward to verify from Equation (15) that an increase in η raises heterogeneity, i.e. $\frac{d\Delta}{d\eta} > 0$. Moreover, by Proposition 1, $\frac{\partial \theta^*}{\partial \Delta} < 0$. Combining these results together, we can see that the effect of η on θ^* is negative. In other words, when assets become more bank-specific, all banks become more stable, including those whose asset performances end up being weaker. Therefore, by (22), increasing η up to $\bar{\eta}$ is unambiguously welfare-improving. Notably, existing literature studying downsides of asset commonality typically argues that a larger degree of asset dispersion is associated with fewer systemic crises but more defaults of weak banks (e.g. Wagner, 2010 and 2011; Ibragimov et al., 2011; Cabrales et al., 2017). Different from these papers that focus on fundamental defaults, we study panic-driven runs and emphasize that more diverse asset performances are associated with lower fragility even for weak banks due to alleviation of cross-bank fire-sale externalities.

Corollary 1. *If $\eta \geq \bar{\eta}$, then cross-bank interactions are trivial, $\Delta \geq \bar{\Delta}$. Run threshold of strong banks $\theta_s^*(\eta)$ decreases in η. Run threshold of weak banks $\theta_w^*(\eta)$ increases in η. The average fragility $\theta^*(\eta)$ does not change with η.*

Proof. See Appendix B.3. □

Proposition 3 and Corollary 1 show that the effect of η on fragility is not "symmetric" in the following sense. If $\eta < \bar{\eta}$, an increase in η leads to a higher bank heterogeneity Δ, lower

Run thresholds $\theta^*(\eta)$, $\theta_s^*(\eta)$, $\theta_w^*(\eta)$

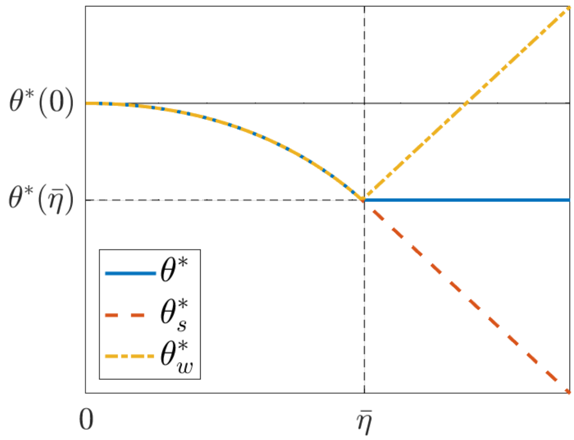

Figure 2: Run thresholds of strong (θ_s^*) and weak (θ_w^*) banks, and their average ($\theta^* = \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$) as functions of the size of bank-specific shock η. Parametrization: $\bar{m} = 0.55$, $\lambda(m) = 1 + m^2$, $F_\epsilon(\cdot)$ is truncated standard normal over $[-1, 1]$, η varies from 0 to 0.05.

fragility for all banks and, hence, higher welfare. If $\eta > \bar{\eta}$, an increase in η has different implications for strong and weak banks. Specifically, if η increases beyond $\bar{\eta}$, banks become so different that cross-bank strategic interactions become trivial, i.e. $\Delta > \bar{\Delta}$. As there are no strategic uncertainties across investors of different banks, the run thresholds of strong and weak banks decouple. That is, the run thresholds θ_s^* and θ_w^*, given by (17) and (18), are not infinitely close and no longer depend on Δ. A further increase in η does not affect cross-bank fire-sale externalities but only further strengthens (weakens) asset performances of strong (weak) banks. As a result, the strong-bank run threshold keeps declining while the weak-bank run threshold starts to increase (see Figure 2). Moreover, since an increase in η does not change the average bank fundamentals, average fragility θ^* remains constant. Overall, our results suggest that the planner should seek to set η to be at least as high as $\bar{\eta}$. Whether welfare (21) increases or decreases in η when $\eta > \bar{\eta}$ depends on the ex-ante distribution of the aggregate fundamental.

Another issue that ring-fencing policies intend to address is a "too-big-to-fail" problem. The following corollary points out that size per se is not the key to financial stability in our model and downsizing banks into identical clones has no impact.[23] Behind the result is that investors of ring-fenced subsidiaries are still interconnected through the asset

[23]Given the focus of this paper is on cross-bank interactions, certain features that are important to study a "too-big-to-fail" issue, such as banks' moral hazard, are absent from our model.

market. The fire-sale complementarity across these subsidiaries resembles the bank-run complementarity across a large group of investors in the merged bank. The objective of ring-fencing, therefore, should not be only to downsize banks' balance sheets but also to achieve an optimal level of asset dispersion across institutions.

Corollary 2. *If ring-fencing does not affect η, then it does not affect fragility.*

Proof. The corollary follows directly from the scale invariance of $\pi(z_i, m_i, m)$. □

4.2 Disclosure of bank-specific information

In this section, we extend the model by allowing for noisy information about bank-specific fundamentals. We then analyze disclosure policies that affect the quality of bank-specific information available to bank investors.

In particular, investor l of bank i receives two distinct noisy signals. The signal about the aggregate fundamental is the same as in the baseline model, $s_{il} = \theta + \sigma\epsilon_{il}$. In addition, she receives a signal d_i about the bank-specific component ζ_i. This signal takes two values, G and B, with a probability mass function specified below.

$$\mathbb{P}(d_i = G|\zeta_i = \eta) = \mathbb{P}\left(d_i = B\,|\zeta_i = -\eta\right) = \alpha \in \left[\frac{1}{2}, 1\right].$$

Parameter α captures the quality of bank-specific information. Under this signal structure, a regulator can vary α by changing the stringency of bank disclosure policy. Below, we discuss how α affects fragility and welfare.

Denote the posterior belief about the probability of bank i being strong as p_G if $d_i = G$ and p_B if $d_i = B$. We have $p_G = \alpha \geq \frac{1}{2} \geq p_B = 1 - \alpha$. The equality holds if and only if $\alpha = \frac{1}{2}$, that is, signals are uninformative about bank-specific fundamentals. Another special case is perfect bank-specific information as in our baseline setting, $\alpha = 1$.

Proposition 4. *The model with imperfect information about bank-specific fundamentals is equivalent to the main model in which bank-specific shocks take values $\eta^{eff}(\alpha)$ with probability $\frac{1}{2}$ and $-\eta^{eff}(\alpha)$ with probability $\frac{1}{2}$, where $\eta^{eff}(\alpha) = (2\alpha - 1)\eta$.*

Proof. See Appendix B.4. □

The main takeaway from Proposition 4 is that besides reshaping asset dispersion directly, regulators can improve stability by manipulating investors' beliefs about it. If signals about bank-specific fundamentals are uninformative, then bank investors perceive strong and weak banks as homogeneous, $\eta^{eff}\left(\frac{1}{2}\right) = 0$. Run decisions of investors of strong and weak banks are then completely synchronized, and the financial system is fragile. If signals are informative, $\alpha > \frac{1}{2}$, investors are able to differentiate between banks. In that case, runs become heterogeneous, $\Delta > 0$. A more stringent disclosure policy improves the quality of bank-specific information α and enlarges the perceived difference between bank asset performances $\eta^{eff}(\alpha)$. In the presence of nontrivial cross-bank strategic interactions, this increases heterogeneity, enhances stability of the financial system by Proposition 3, and improves welfare given by (22). In particular, the planner should always seek to set α such that $\eta^{eff}(\alpha)$ is not smaller than $\bar{\eta}$ that is given by (23).

Existing literature on disclosure policies highlights various costs and benefits of disclosing bank-specific information. We highlight a novel benefit of disclosing bank-specific information: It enhances financial stability by boosting bank heterogeneity. Notably, some papers argue that disclosure should be state-dependent so that bank information is opaque in good times and transparent in bad times (e.g. Bouvard et al., 2015; Goldstein and Leitner, 2018). Specifically, disclosing bank-specific information in bad times induces weak banks to fail but, at the same time, saves strong banks. In our model, disclosing bank-specific information makes weak-bank investors aware of the weak fundamentals of their banks but, at the same time, less worried about spillovers from strong banks. The latter effect dominates in the presence of reinforcing complementarities, and weak-bank investors' run incentives weaken. As a result, instead of inducing runs, disclosure reduces fragility of even weak banks.

Naturally, in the absence of cross-bank strategic uncertainty, disclosure can no longer indirectly stabilize weak banks because fire-sale pressure from strong banks is already at its minimum level. Specifically, when $\eta^{eff}(\alpha) > \bar{\eta}$, cross-bank strategic interactions are trivial, and disclosing more bank-specific information stabilizes strong banks, destabilizes weak banks and has no impact on the average fragility (see Figure 2). In this case, whether additional disclosure is desirable depends on the ex-ante distribution of the aggregate fundamental.

4.3 Asset market interventions

One of the distress resolution approaches that regulators frequently turn to during economic downturns is a liquidity injection into asset markets. Prominent examples include the U.S. government's purchases of distressed assets during the 2007–2008 financial crisis and corporate bonds and bond ETFs via the Secondary Market Corporate Credit Facility during the COVID-19 pandemic. In our framework, such policies reduce fire-sale discounts. Keeping bank heterogeneity fixed, a reduction in fire-sale discounts stabilizes banks directly. As we show below, such policies also impact bank heterogeneity and thus affect fragility indirectly.[24]

We consider an extension of our baseline model in which the regulator injects $L \in [0, \bar{m})$ into the asset market. Specifically, asset prices remain at their fundamental levels if the aggregate liquidity needs do not exceed L. Any liquidity needs beyond this point are fulfilled by outside investors as in the baseline model. As a result, the fire-sale discount factor in the asset market is

$$\hat{\lambda}(m, L) = \begin{cases} 1 & \text{if } m < L, \\ \lambda\left(m - L\right) & \text{if } m \geq L, \end{cases}$$

where $\lambda(\cdot)$ is the fire-sale discount factor in the baseline model and m is the total mass of runners in the economy. To evaluate the impacts of liquidity injections, we establish comparative statics with respect to L in Proposition 5.

Proposition 5. *Suppose that $\lambda''(\cdot) \geq 0$ and $\eta \in (0, \bar{\eta})$, where $\bar{\eta}$ is defined in Equation (23). There exists a decreasing function $\bar{L}(\eta) > 0$ such that if $L \in \left(0, \bar{L}\left(\eta\right)\right)$, then cross-bank interactions are nontrivial, $\Delta < \bar{\Delta}$, and run thresholds of strong and weak banks are infinitely close to θ^*. Moreover, $\frac{d\theta^*}{dL} = \frac{\partial \theta^*}{\partial L} + \frac{\partial \theta^*}{\partial \Delta}\frac{d\Delta}{dL}$, where*

1. *The direct effect is stabilizing, $\frac{\partial \theta^*}{\partial L} < 0$;*

[24] Regulators often act as lenders of last resort (LOLR) and inject liquidity to financial institutions. Similar to a liquidity injection into the asset market, a LOLR intervention stabilizes the financial system both directly and indirectly.

2. The indirect effect is stabilizing, $\frac{\partial \theta^*}{\partial \Delta} < 0$ *and* $\frac{d\Delta}{dL} > 0$*.*

Proof. See Appendix B.5. $\square$

Different from ring-fencing and disclosure policies that influence fragility only through their impacts on bank heterogeneity, a liquidity injection has a direct stabilizing effect, captured by $\frac{\partial \theta^*}{\partial L} < 0$. Importantly, even though liquidity injections are broad-based—that is, the regulator does not restrict the purchase to assets owned by a particular group of banks—strong and weak banks are affected differently. The reason is that strong-bank investors expect to receive a higher implicit subsidy *in the event of a run.* In our model, what determines the run threshold for a given bank is the belief of this bank's marginal investor. Given strong banks receive a positive bank-specific shock, a strong-bank investor is on the margin of running when the market liquidity conditions are particularly dire. In contrast, a weak-bank investor is on the margin of running under milder fire-sale pressure. Formally, from Equation (9), if a strong-bank marginal investor expects a fraction x of "non-sleepy" strong-bank investors to run, then she expects a larger fraction $F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) > x$ of "non-sleepy" weak-bank investors to run. In contrast, from Equation (10), a weak-bank marginal investor expecting a fraction x of "non-sleepy" weak-bank investors to run expects a smaller fraction $F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) < x$ of "non-sleepy" strong-bank investors to run. Therefore, inefficient liquidations, occurring when the total mass of runs in the economy exceeds the liquidity injection size L, are more likely from the marginal strong-bank investor's perspective than from the marginal weak-bank investor's perspective. Consequently, an increase in the liquidity injection size reduces fire-sale pressure on strong banks in a larger number of states, which implies $\frac{\partial I_s}{\partial L} < \frac{\partial I_w}{\partial L} < 0$.[25] This further strengthens strong banks relative to weak banks, thereby increasing bank heterogeneity, $\frac{d\Delta}{dL} > 0$, and stabilizing the financial system indirectly, $\frac{\partial \theta^*}{\partial \Delta} < 0$.

For completeness, the corollary below summarizes the case with trivial cross-bank interactions. In this case, heterogeneity ceases to have an effect and liquidity injection

[25]Although this effect arises even if the fire-sale discount function $\lambda(\cdot)$ is linear, it becomes more pronounced if it is convex. In that case, an increase in L is particularly helpful for strong banks because their marginal investors expect a large fire-sale discount on average.

functions only through the direct effect.

Corollary 3. *If $\eta \geq \bar{\eta}$ or $L \geq \bar{L}(\eta)$, cross-bank interactions are trivial, $\Delta \geq \bar{\Delta}$. Liquidity injection reduces the average fragility and fragilities of strong and weak banks through the direct effect only,* $\frac{d\theta^*}{dL} = \frac{\partial \theta^*}{\partial L} < 0$, $\frac{d\theta^*_s}{dL} = \frac{\partial \theta^*_s}{\partial L} < 0$, $\frac{d\theta^*_w}{dL} = \frac{\partial \theta^*_w}{\partial L} < 0$.

Proof. See Appendix B.5. □

Given that liquidity injections can stabilize the financial system both directly and indirectly, it is important to understand how the latter effect compares to the former quantitatively. We conduct a simple numerical exercise in the context of the Great Recession. In particular, we pick the size of bank-specific shocks $\eta = 0.025$ to match the cross-sectional dispersion in annual returns on assets of U.S. banks during the 2007–2008 financial crisis (our data is from the FR Y-9C filings; see Appendix A for more details). We set $\bar{m} = 0.57$, corresponding to the fraction of bank liabilities that is not covered by deposit insurance in 2007Q4 and is thus subject to runs. We parametrize $\lambda(m - L) = 1 + (m - L)^2$. Under such a choice, when a systemic bank run occurs in the absence of any government liquidity injections, that is, $m = \bar{m}$ and $L = 0$, asset prices are 25% below their fundamental values. This value is in line with the estimates of James (1991) and Granja, Matvos, and Seru (2017) who document comparable discounts in prices paid for assets of failed banks during the Savings and Loan Crisis and the 2007–2008 financial crisis, respectively.[26] Finally, we choose the uniform noise distribution with a support $[-1, 1]$. Our results are not sensitive to this choice.

We normalize the aggregate amount of bank assets to one and vary the liquidity injection size L from 0 to 0.01. Our parametrization implies nontrivial cross-bank interactions for all $L \in [0, 0.01]$. We investigate how the common run threshold $\theta^* = \theta^*(L)$ changes with the size of liquidity injection in two scenarios. In the baseline case, bank heterogeneity $\Delta = \Delta(L)$ adjusts as L increases such that $\theta^*(L)$ changes both due to direct and indirect effects. This case is illustrated by the blue solid line in the left panel of Figure 3.

[26] Our parametrization of the fire-sale discount function is conservative. For example, during the 2007–2008 crisis, the government actively intervened in the financial sector, likely having prevented more widespread runs. In the absence of such interventions, a collapse in asset prices might have been much more substantial.

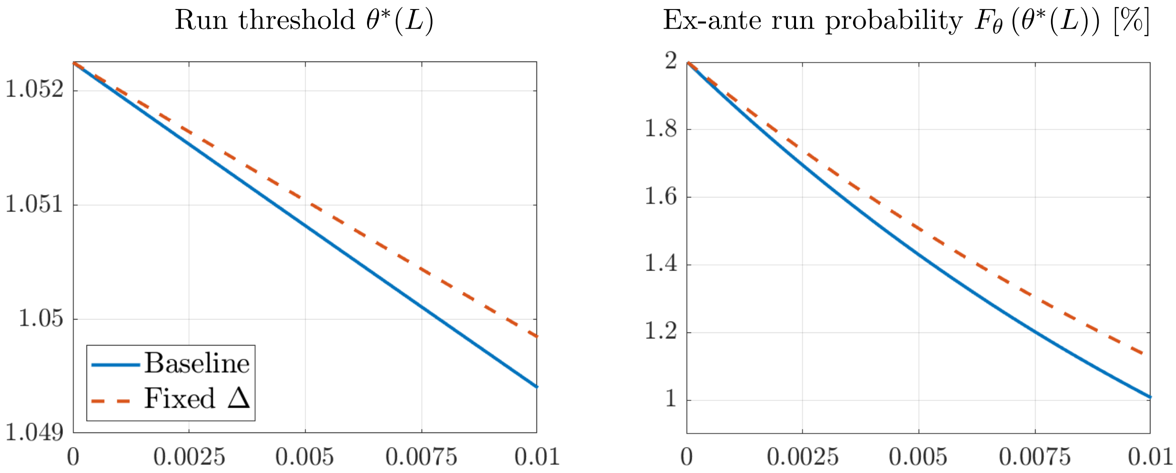

Figure 3: Run threshold θ^* (left panel) and ex-ante probability of a systemic crisis $F_\theta(\theta^*)$ (right panel; in percentage) as functions of the liquidity injection size L. In the "Baseline" case, heterogeneity $\Delta = \Delta(L)$ adjusts as L changes. In the "Fixed Δ" case, bank heterogeneity is fixed at its initial level $\Delta(0)$. See text for parametrization.

The red dashed line shows how the run threshold changes with L if bank heterogeneity is held fixed, $\Delta = \Delta(0)$. Therefore, it reflects only the direct effect. The difference suggests that the indirect effect accounts for about 15% of the total reduction in the average run threshold as L changes from 0 to 0.01.

To map the change in the run threshold to that in the probability of a crisis, we need to calibrate the prior distribution of the aggregate fundamental $F_\theta(\cdot)$. We assume that the prior distribution of θ is log-normal with parameters μ_θ and σ_θ. Using the FR Y-9C filings, we calibrate $\sigma_\theta = 0.01$ to match the standard deviation of the average returns on assets for U.S. banks in the pre-crisis years of 1991–2006. We pick $\mu_\theta = 0.0715$ so that the annual frequency of systemic crises in the model is $F_\theta(\theta^*) = 2\%$, a number consistent with the historical data (e.g. Jordà, Schularick, and Taylor, 2017; Romer and Romer, 2017). The right panel of Figure 3 shows how the ex-ante run probability $F_\theta(\theta^*(L))$ changes with the size of liquidity injection when bank heterogeneity adjusts with L (solid blue line) and when bank heterogeneity is held fixed at $\Delta = \Delta(0)$ (red dashed line). We observe that an increase in L from 0 to 0.01 implies a significant decline in the ex-ante run probability from 2% to 1%. The indirect effect accounts for a nontrivial part of this decline: If bank heterogeneity is held fixed, the run probability declines from 2% to 1.13%.

4.4 Liquidity buffers

In our baseline model, we assume that banks invest all capital they raise from investors in long-term assets. As a result, if banks face runs at $t = 1$, they have to liquidate some of their long-term assets. Furthermore, banks are unable to acquire each other's assets. In reality, stronger banks frequently acquire assets of failing banks. This is often seen as a desirable outcome because banks are likely more efficient users of those assets than outside investors. In this section, we extend our model by allowing banks to hold a fraction $l \in [0, \bar{m})$ of their capital in cash that has a unit return. We treat l as an exogenous parameter and study how changes in l affect bank heterogeneity and fragility. This exercise can be seen as investigating the impacts of changing liquidity requirements on the economy. In Appendix C.7, we solve for an endogenous l and discuss why banks underinvest in liquid assets relative to the social optimum.

More specifically, suppose that bank i invests a fraction l of its assets in cash and the remaining $1-l$ fraction in a long-term project. If the mass of early withdrawals is $m_i > l$, the bank has to liquidate some of its long-term assets. If, on the contrary, $m_i < l$, then the bank can repay runners without liquidating its long-term assets. Furthermore, it can acquire assets liquidated by other banks (if any). We assume that banks are efficient users of liquidated assets. As such, bank i acquiring one unit of bank j's long-term assets expects to get a gross return z_j. Therefore, the price bank i is willing to pay for each unit of bank j's assets is z_j.

In this extended setup, strong banks acquire assets of weak banks in some states of the world. If a marginal weak-bank investor believes that a fraction x of "non-sleepy" weak-bank investors run, then her perceived mass of runners in strong banks is $\bar{m}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) < \bar{m}x$. Thus, even if a marginal weak-bank investor expects that her bank is unable to fully repay runners using its liquidity buffer, i.e. $\bar{m}x > l$, then strong banks can step in and acquire at least part of liquidated assets. In particular, if $\bar{m}x < l < \bar{m}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)$, weak banks do not face any inefficient liquidations because strong banks fully satisfy weak banks' liquidity needs. If $\bar{m}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) > l$, strong banks are unable to acquire all assets liquidated by weak banks. Still, they alleviate fire-sale pressure on weak banks by reducing the amount of assets outside investors have to absorb.

In contrast, weak banks never acquire assets of strong banks. If a marginal strong-bank investor believes that a fraction x of "non-sleepy" strong-bank investors run, then her perceived mass of runners in weak banks is $\bar{m}F_{\epsilon}\left(\Delta + F_{\epsilon}^{-1}\left(x\right)\right) > \bar{m}x$. Therefore, if a marginal strong-bank investor expects that her bank is unable to fully repay runners using its liquidity buffer, i.e. $\bar{m}x > l$, then weak banks are also unable to do so, and liquidated assets are absorbed by inefficient outside investors. Nevertheless, liquidity buffers held by weak banks reduce fire-sale pressure on strong banks because weak banks can cover part of early withdrawals with their internal funds and thus sell fewer assets to outside investors. This has a large beneficial effect on strong banks because they experience runs when market liquidity conditions are particularly dire.

The following proposition establishes how changes in the liquidity buffer size affect financial fragility.

Proposition 6. *Suppose that $\lambda''(\cdot) \geq 0$ and $\eta \in (0, \bar{\eta})$, where $\bar{\eta}$ is defined in Equation (23). There exists a decreasing function $\bar{l}(\eta) > 0$ such that if $l \in \left(0, \bar{l}\left(\eta\right)\right)$, then cross-bank interactions are nontrivial, $\Delta < \bar{\Delta}$, and run thresholds of strong and weak banks are infinitely close to θ^*. Moreover, $\frac{d\theta^*}{dl} = \frac{\partial\theta^*}{\partial l} + \frac{\partial\theta^*}{\partial\Delta}\frac{d\Delta}{dl}$, where*

1. *The direct effect is stabilizing, $\frac{\partial\theta^*}{\partial l} < 0$;*

2. *The indirect effect is stabilizing, $\frac{\partial\theta^*}{\partial\Delta} < 0$ and $\frac{d\Delta}{dl} > 0$.*

Proof. See Appendix B.6. □

Proposition 6 shows that the impact of an increase in the liquidity buffer size l on fragility is qualitatively analogous to that of an increase in the liquidity injection size L discussed in Section 4.3. Naturally, a higher l stabilizes banks directly, $\frac{\partial\theta^*}{\partial l} < 0$, because it implies that banks have to liquidate fewer long-term assets if runs occur. A higher l also stabilizes banks indirectly. Similar to the intuition behind Proposition 5, a strong-bank investor is on the margin of running under worse liquidity conditions than a weak-bank marginal investor. Therefore, an increase in l has a stronger positive effect on a strong-bank marginal investor than on a weak-bank marginal investor, leading to a higher bank heterogeneity, $\frac{\partial \Delta}{\partial l} > 0$.

For completeness, the corollary below summarizes the case with trivial cross-bank interactions where heterogeneity ceases to have an effect.

Corollary 4. *If $\eta \geq \bar{\eta}$ or $l \geq \bar{l}(\eta)$, cross-bank interactions are trivial, $\Delta \geq \bar{\Delta}$. Liquidity buffer reduces the average fragility and fragilities of strong and weak banks through the direct effect only, $\frac{d\theta^*}{dl} = \frac{\partial \theta^*}{\partial l} < 0$, $\frac{d\theta^*_s}{dl} = \frac{\partial \theta^*_s}{\partial l} < 0$, $\frac{d\theta^*_w}{dl} = \frac{\partial \theta^*_w}{\partial l} < 0$.*

Proof. See Appendix B.6. □

5 Concluding remarks

This paper analyzes interactions between fragile banks through asset fire sales. Such interactions can lead to a spread of panics across institutions and thus a systemic crisis. The key feature of our model is the reinforcement between within- and cross-bank complementarities, that is, bank runs and fire sales. We highlight the key factor—bank heterogeneity—that governs cross-bank interactions and plays an important role in determining the fragility of the entire financial system.

The broad message of our policy analyses is that in a financial system featuring interconnected fragilities, various factors can affect financial stability through changing bank heterogeneity. Regulators should take this indirect effect into account when monitoring bank activities and developing intervention strategies. Without doing so, policies may undermine financial stability by making runs more synchronized.

For instance, regulatory policies that narrow the dispersion in bank asset performances can be harmful for stability. One example is mergers and acquisitions in the banking system. Another example is direct financial support—or a belief about it shared by market participants—biased toward weak banks. Although such financial support directly enhances weak banks' resilience to negative shocks, its effectiveness can be hindered by the resulting commonality in (perceived) asset performances. Furthermore, when banks face multiple capital regulations, regulatory arbitrage incentive pushes banks to converge in their business models and asset holdings (Greenwood, Stein, Hanson, and Sunderam, 2017).

Stress tests is another widely-used policy tool that can lead to an unintended reduction in heterogeneity. Currently, regulators test bank resilience against a common set of stress scenarios. Banks that are more likely to fail these tests—weak banks in our paper—are required by regulators to increase their liquidity buffers or equity cushions, which can raise their similarity to strong banks. Furthermore, anticipating such corrective actions, weak banks might act preemptively and strengthen their balance sheets against stress test scenarios, again making the financial system more homogeneous. As a result of higher similarity across banks, heterogeneity diminishes, which can undermine the effectiveness of stress tests. Our paper suggests that a certain difference between banks' resistance to a common set of stress scenarios is desirable.

Relatedly, regulators should closely monitor aggregate market conditions and conduct interventions in a timely manner. Recall that when strong banks are forced to liquidate their assets, the asset market is already highly stressed due to massive liquidations by weak banks. If regulators commit to intervene in situations when fire-sale discounts are especially high, then this can be particularly helpful to strong banks. As our results in Section 4.3 suggest, this can have a sizeable stabilizing effect through enlarging heterogeneity and alleviating the vicious loop between bank runs and deteriorating liquidity conditions.

Overall, various bank activities and regulatory policies might affect heterogeneity and, thus, impact financial stability. Regulators should take a holistic approach and combine different regulatory tools to maintain a certain level of heterogeneity in the financial system.

References

Acharya, V. V. (2009): "A Theory of Systemic Risk and Design of Prudential Bank Regulation," *Journal of Financial Stability*, 5, 224–255.

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. Richardson (2017):

"Measuring Systemic Risk," *Review of Financial Studies*, 30, 2–47.

Adrian, T. and M. K. Brunnermeier (2016): "CoVaR," *American Economic Review*, 106, 1705–1741.

Allen, F., A. Babus, and E. Carletti (2012): "Asset Commonality, Debt Maturity and Systemic Risk," *Journal of Financial Economics*, 104, 519–534.

Allen, F. and D. Gale (1994): "Limited Market Participation and Volatility of Asset Prices," *American Economic Review*, 84, 933–955.

Angeletos, G.-M. and I. Werning (2006): "Crises and Prices: Information Aggregation, Multiplicity, and Volatility," *American Economic Review*, 96, 1720–1736.

Athey, S. (2002): "Monotone Comparative Statics under Uncertainty," *Quarterly Journal of Economics*, 117, 187–223.

Bouvard, M., P. Chaigneau, and A. D. Motta (2015): "Transparency in the Financial System: Rollover Risk and Crises," *Journal of Finance*, 70, 1805–1837.

Brownlees, C. and R. F. Engle (2017): "SRISK: A Conditional Capital Shortfall Measure of Systemic Risk," *Review of Financial Studies*, 30, 48–79.

Cabrales, A., P. Gottardi, and F. Vega-Redondo (2017): "Risk Sharing and Contagion in Networks," *Review of Financial Studies*, 30, 3086–3127.

Carletti, E., I. Goldstein, and A. Leonello (2020): "The Interdependence of Bank Capital and Liquidity," *BAFFI CAREFIN Centre Research Paper*.

Carlsson, H. and E. van Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.

Chen, Q., I. Goldstein, and W. Jiang (2010): "Payoff Complementarities and Financial Fragility: Evidence from Mutual Fund Outflows," *Journal of Financial Economics*, 97, 239–262.

Choi, D. B. (2014): "Heterogeneity and Stability: Bolster the Strong, Not the Weak," *Review of Financial Studies*, 27, 1830–1867.

Cifuentes, R., G. Ferrucci, and H. S. Shin (2005): "Liquidity Risk and Contagion," *Journal of the European Economic Association*, 3, 556–566.

Cong, L. W., S. R. Grenadier, and Y. Hu (2020): "Dynamic Interventions and Informational Linkages," *Journal of Financial Economics*, 135, 1–15.

Dai, L., D. Luo, and M. Yang (2021): "Disclosure of Bank-specific Information and Stability of Financial Systems," *Working paper.*

Diamond, D. W. and P. H. Dybvig (1983): "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 91, 401–419.

Diamond, D. W. and R. G. Rajan (2005): "Liquidity Shortages and Banking Crises," *Journal of Finance*, 60, 615–647.

Eisenbach, T. M. (2017): "Rollover Risk as Market Discipline: A Two-Sided Inefficiency," *Journal of Financial Economics*, 126, 252–269.

Elsinger, H., A. Lehar, and M. Summer (2006): "Risk Assessment for Banking Systems," *Management Science*, 52, 1301–1314.

Enria, A. (2018): "Fragmentation in Banking Markets: Crisis Legacy and the Challenge of Brexit," *Speech delivered at BCBS-FSI High Level Meetign for Europe on Banking Supervision.*

Frankel, D. M., S. Morris, and A. Pauzner (2003): "Equilibrium Selection in Global Games with Strategic Complementarities," *Journal of Economic Theory*, 108, 1–44.

Goldstein, I. (2005): "Strategic Complementarities and the Twin Crises," *Economic Journal*, 115, 368–390.

GOLDSTEIN, I., H. JIANG, AND D. T. NG (2017): "Investor Flows and Fragility in Corporate Bond Funds," *Journal of Financial Economics*, 126, 592–613.

GOLDSTEIN, I. AND Y. LEITNER (2018): "Stress Tests and Information Disclosure," *Journal of Economic Theory*, 177, 34–69.

GOLDSTEIN, I. AND A. PAUZNER (2005): "Demand-Deposit Contracts and the Probability of Bank Runs," *Journal of Finance*, 60, 1293–1327.

GOMES, J. F. (2001): "Financing Investment," *American Economic Review*, 91, 1263–1285.

GRANJA, J., G. MATVOS, AND A. SERU (2017): "Selling Failed Banks," *Journal of Finance*, 72, 1723–1784.

GREENWALD, B. C. AND J. E. STIGLITZ (1986): "Externalities in Economies with Imperfect Information and Incomplete Markets," *Quarterly Journal of Economics*, 101, 229–264.

GREENWOOD, R., J. C. STEIN, S. G. HANSON, AND A. SUNDERAM (2017): "Strengthening and Streamlining Bank Capital Regulation," *Brookings Papers on Economic Activity*, 2017, 479–565.

HALDANE, A. G. (2009): "Rethinking the Financial Network," *Speech delivered at the Financial Student Association, Amsterdam.*

HELLWIG, C., A. MUKHERJI, AND A. TSYVINSKI (2006): "Self-fulfilling Currency Crises: The Role of Interest Rates," *American Economic Review*, 96, 1769–1787.

HUANG, X., H. ZHOU, AND H. ZHU (2009): "A Framework for Assessing the Systemic Risk of Major Financial Institutions," *Journal of Banking & Finance*, 33, 2036–2049.

IBRAGIMOV, R., D. JAFFEE, AND J. WALDEN (2011): "Diversification Disasters," *Journal of Financial Economics*, 99, 333–348.

James, C. (1991): "The Losses Realized in Bank Failures," *Journal of Finance*, 46, 1223–1242.

Jordà, Ò., M. Schularick, and A. M. Taylor (2017): "Macrofinancial History and the New Business Cycle Facts," *NBER Macroeconomics Annual*, 31, 213–263.

Kiyotaki, N. and J. Moore (1997): "Credit Cycles," *Journal of Political Economy*, 105, 211–248.

Kopytov, A. (2023): "Booms, Busts, and Common Risk Exposures," *Journal of Finance*, 78, 3299–3341.

Kreicher, L. L. and R. N. McCauley (2018): "The New US Intermediate Holding Companies: Reducing or Shifting Assets?" *BIS Quarterly Review*.

Kuong, J. C.-F. (2021): "Self-Fulfilling Fire Sales: Fragility of Collateralized Short-Term Debt Markets," *Review of Financial Studies*, 34, 2910–2948.

Liu, X. (2016): "Interbank Market Freezes and Creditor Runs," *Review of Financial Studies*, 29, 1860–1910.

——— (2018): "Diversification and Systemic Bank Runs," *Working Paper*.

Lorenzoni, G. (2008): "Inefficient Credit Booms," *Review of Economic Studies*, 75, 809–833.

Morris, S. and H. S. Shin (1998): "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *American Economic Review*, 88, 587–597.

——— (2003): *Global Games: Theory and Applications*, Advances in Economics and Econometrics: Theory and Applications, 56–114, Econometric Society Monographs.

Parlatore, C. (2015): "Transparency and Bank Runs," *Working paper*.

——— (2016): "Fragility in Money Market Funds: Sponsor Support and Regulation," *Journal of Financial Economics*, 121, 595–623.

ROCHET, J.-C. AND X. VIVES (2004): "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?" *Journal of the European Economic Association*, 2, 1116–1147.

ROMER, C. D. AND D. H. ROMER (2017): "New Evidence on the Aftermath of Financial Crises in Advanced Countries," *American Economic Review*, 107, 3072–3118.

SÁKOVICS, J. AND J. STEINER (2012): "Who Matters in Coordination Problems?" *American Economic Review*, 102, 3439–3461.

SHAFFER, S. (1994): "Pooling Intensifies Joint Failure Risk," *Research in Financial Services*, 6, 249–280.

SHEN, L. AND J. ZOU (2020): "Intervention with Screening in Panic-Based Runs," *Available at SSRN 3137172*.

SHLEIFER, A. AND R. W. VISHNY (1992): "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," *Journal of Finance*, 47, 1343–1366.

SONG, F. AND A. V. THAKOR (2022): "Market Freeze and Bank Capital Structure Heterogeneity," *Management Science*, Forthcoming.

STIGLITZ, J. E. (2010): "Risk and Global Economic Architecture: Why Full Financial Integration May be Undesirable," *American Economic Review: Papers and Proceedings*, 100.

UHLIG, H. (2010): "A Model of a Systemic Bank Run," *Journal of Monetary Economics*, 57, 78–96.

UPPER, C. (2011): "Simulation Methods to Assess the Danger of Contagion in Interbank Markets," *Journal of Financial Stability*, 7, 111–125.

WAGNER, W. (2010): "Diversification at Financial Institutions and Systemic Crises," *Journal of Financial Intermediation*, 19, 373–386.

——— (2011): "Systemic Liquidation Risk and the Diversity-Diversification Trade-Off," *Journal of Finance*, 66, 1141–1175.

Yellen, J. (2013): "Interconnectedness and Systemic Risk: Lessons from the Financial Crisis and Policy Implications," *Speech delivered at the AEA/AFA Joint Luncheon, San Diego*.

A Data appendix for Section 4.3

Our data on U.S. bank holding companies is from the FR Y-9C filings and covers the period between 1991 and 2018. Each year, we focus on 100 largest parent-level bank holding companies to reduce the role of outliers; our results are not sensitive to this cutoff. Each year t, we compute returns on assets R_{it} for each bank i as the sum of interest and nontinterest income net of noninterest expense, divided by the last-year total assets. In the model, θ is the aggregate productivity of bank assets before payments to investors and in the absence of any inefficient liquidations. Therefore, to estimate σ_θ, we compute time-series standard deviation of the cross-bank average of R_{it} in the years before the 2007–2008 financial crisis. The parameter η governs the cross-sectional standard deviation in productivities of bank assets. Since this standard deviation can be time-varying and since our focus is on understanding investors' run decisions, we pick η to match the cross-sectional standard deviation in R_{it} during the 2007–2008 period.

B Proofs

B.1 Liquidation price

This appendix proves Lemma 1.

Proof. Outside investors solve the problem

$$\max_{\{k_i\}_{i\in[0,1]}} \int z_i k_i di - g\left(\int p_i k_i di\right).$$

The first-order condition implies that the liquidation price of bank i's assets p_i is proportional to its fundamental z_i,

$$p_i = \frac{z_i}{g'(L)} \quad \forall i \in [0,1], \tag{24}$$

where $L \equiv \int p_i k_i di$. Aggregation of the market clearing conditions for individual banks implies that the total liquidity demand m equals to the liquidity supply L:

$$m_i = p_i k_i \Rightarrow m = \int p_i k_i di = L.$$

Therefore, we obtain the equilibrium asset prices

$$p_i(z_i, m) = \frac{z_i}{\lambda(m)},$$

where $\lambda(m) \equiv g'(m)$. Since $\lambda'(m) = g''(m) > 0$, the liquidation price $p_i(z_i, m)$ is a decreasing function of the total mass of runners m for any $i \in [0, 1]$. $\square$

B.2 Role of two complementarities

We prove Proposition 2. Note that Proposition 1 is a special case of Proposition 2 with $\pi_1(m_i) = \frac{1}{1-m_i}$ and $\pi_2(m_i, m) = -\frac{\lambda(m)m_i}{1-m_i}$. In Appendix C.2, we explore a general net payoff function $\pi(z_i, m_i, m)$ without imposing a specific functional form.

Proof. If $\Delta \equiv \lim_{\sigma\to 0} \frac{\theta_w^* - \theta_s^*}{\sigma} < \bar{\Delta}$, the model features nontrivial cross-bank strategic interactions. The analogues of Equations (7) and (8) from the main text are

$$\int_0^1 \left[(\theta_s^* + \eta)\pi_1(\bar{m}x) + \pi_2\left(\bar{m}x, \frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right)\right)\right] dx = 0, \tag{25}$$

$$\int_0^1 \left[(\theta_w^* - \eta)\pi_1(\bar{m}x) + \pi_2\left(\bar{m}x, \frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\right)\right] dx = 0. \tag{26}$$

Since we focus on the limiting case with vanishing information friction, the two run thresholds are infinitely close to the average threshold $\theta^* \equiv \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$.

Differentiating (25) and (26) with respect to Δ, we can derive

$$\frac{\partial\theta_s^*}{\partial\Delta} = -\frac{\frac{\bar{m}}{2}\int_0^{F_\epsilon(\bar{\epsilon}-\Delta)} \frac{\partial}{\partial m}\left\{\pi_2\left(\bar{m}x, \frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right)\right)\right\} f_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) dx}{\int_0^1 \pi_1\left(\bar{m}x\right) dx}, \tag{27}$$

$$\frac{\partial\theta_w^*}{\partial\Delta} = \frac{\frac{\bar{m}}{2}\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \frac{\partial}{\partial m}\left\{\pi_2\left(\bar{m}x, \frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\right)\right\} f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) dx}{\int_0^1 \pi_1\left(\bar{m}x\right) dx} =$$

$$\frac{\frac{\bar{m}}{2}\int_0^{F_\epsilon(\bar{\epsilon}-\Delta)} \frac{\partial}{\partial m}\left\{\pi_2\left(\bar{m}F_\epsilon\left(\Delta + F_\epsilon^{-1}(y)\right), \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(y)\right) + \frac{\bar{m}}{2}y\right)\right\} f_\epsilon\left(\Delta + F_\epsilon^{-1}(y)\right) dy}{\int_0^1 \pi_1\left(\bar{m}x\right) dx}, \tag{28}$$

where (28) is obtained by changing the variable of integration $x \to F_\epsilon\left(\Delta + F_\epsilon^{-1}(y)\right)$, which implies, in particular, $dx = dF_\epsilon\left(\Delta + F_\epsilon^{-1}(y)\right) = \frac{f_\epsilon(\Delta + F_\epsilon^{-1}(y))}{f_\epsilon\left(F_\epsilon^{-1}(y)\right)} dy$.

The impact of bank heterogeneity Δ on θ^* is

$$\frac{\partial\theta^*}{\partial\Delta} = \frac{1}{2}\left(\frac{\partial\theta_s^*}{\partial\Delta} + \frac{\partial\theta_w^*}{\partial\Delta}\right).$$

Comparing (27) and (28), we can see that

$$\frac{\partial^2 \pi}{\partial m \partial m_i} = \frac{\partial^2 \pi_2}{\partial m \partial m_i} \lesseqqgtr 0 \Rightarrow -\frac{\partial \theta_w^*}{\partial \Delta} \gtreqqless \frac{\partial \theta_s^*}{\partial \Delta} \Rightarrow \frac{\partial \theta^*}{\partial \Delta} \lesseqqgtr 0.$$

□

B.3 Ring-fencing

This appendix proves Proposition 3 and Corollary 1.

B.3.1 Proposition 3

Proof. Recall that when cross-bank strategic interactions are nontrivial ($\Delta < \bar{\Delta}$), bank heterogeneity Δ is implicitly defined by Equation (15), which we repeat below,

$$\eta = \frac{1}{2\int_0^1 \frac{dx}{1-\bar{m}x}} \left(I_s(\Delta) - I_w(\Delta) \right). \tag{29}$$

From Equations (9) and (10), it is clear that $I_s(\Delta)$ increases in Δ and $I_w(\Delta)$ decreases in Δ. Hence, $\Delta(\eta)$ is an increasing function. Moreover, $\Delta(0) = 0$ and $\Delta(\bar{\eta}) = \bar{\Delta}$.

Therefore, when $\eta \in (0, \bar{\eta})$, we have $\Delta \in (0, \bar{\Delta})$, and cross-bank strategic interactions are nontrivial. From Equation (14), $\theta^*(\eta) = \theta^*(\Delta(\eta))$. Since $\Delta(\eta)$ increases in η on $(0, \bar{\eta})$, by Proposition 1, run thresholds of strong and week banks are infinitely close to $\theta^*(\eta)$ which decreases in η on $(0, \bar{\eta})$. □

B.3.2 Corollary 1

Proof. If $\eta \geq \bar{\eta}$, Equation (29) does not have a solution $\Delta = \Delta(\eta)$. Therefore, run thresholds θ_s^* and θ_w^* cannot be infinitely close to each other. Then they are given by (17) and (18), respectively. From these equations, it is clear that $\theta_s^*(\eta)$ is a decreasing function, $\theta_w^*(\eta)$ is an increasing function, and $\frac{1}{2}\theta_s^*(\eta) + \frac{1}{2}\theta_w^*(\eta)$ does not depend on η. □

B.4 Noisy information about bank-specific fundamentals

We prove Proposition 4 in this appendix.

Proof. Fraction α of strong-bank investors and fraction $1-\alpha$ of weak-bank investors receive signal G. Therefore, fraction $\frac{1}{2}\alpha + \frac{1}{2}(1-\alpha) = \frac{1}{2}$ of all investors receive this signal. These investors run if their signals about the aggregate fundamental are below

θ_G^*. Similarly, fraction $\frac{1}{2}(1-\alpha)+\frac{1}{2}\alpha=\frac{1}{2}$ of all investors receive signal B. These investors run if their signals about the aggregate fundamental are below θ_B^*. The run thresholds for investors receiving signals G and B are determined by

$$\theta_G^* + p_G\eta - (1-p_G)\eta = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1+\int_0^1 \frac{\lambda(m_G(x,\Delta))\bar{m}x}{1-\bar{m}x}dx\right),$$
$$\theta_B^* + p_B\eta - (1-p_B)\eta = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1+\int_0^1 \frac{\lambda(m_B(x,\Delta))\bar{m}x}{1-\bar{m}x}dx\right),$$

where $\Delta = \lim_{\sigma\to 0}\frac{\theta_B^*-\theta_G^*}{\sigma}$ and

$$m_G(x,\Delta) = \frac{1}{2}\bar{m}x + \frac{1}{2}\bar{m}F_\epsilon(\Delta + F_\epsilon^{-1}(x)),$$
$$m_B(x,\Delta) = \frac{1}{2}\bar{m}F_\epsilon(-\Delta + F_\epsilon^{-1}(x)) + \frac{1}{2}\bar{m}x.$$

Define

$$\eta^{eff}(\alpha) = p_G\eta - (1-p_G)\eta = (2\alpha-1)\,\eta.$$

It straightforward to see that the model described in Section 4.2 boils down to our baseline setting with redefined η. Therefore, the results of Section 3.2 generalize to the case of noisy bank-specific signals. $\square$

B.5 Asset market interventions

This appendix proves Proposition 5 and Corollary 3. Throughout this appendix, we make the following assumption.

Assumption 1. *$\lambda(\cdot)$ is a weakly convex function, that is, $\lambda''(\cdot)\geq 0$.*

B.5.1 Proposition 5

With the modified fire-sale discount function $\hat{\lambda}(m,L)$, we can express the fire-sale pressure terms as

$$\hat{I}_s(\Delta,L) = \int_{x_s(\Delta,L)}^1 \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - L\right)\frac{\bar{m}x}{1-\bar{m}x}dx + \int_0^{x_s(\Delta,L)}\frac{\bar{m}x}{1-\bar{m}x}dx, \quad (30)$$

$$\hat{I}_w(\Delta,L) = \int_{x_w(\Delta,L)}^1 \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - L\right)\frac{\bar{m}x}{1-\bar{m}x}dx + \int_0^{x_w(\Delta,L)}\frac{\bar{m}x}{1-\bar{m}x}dx, \quad (31)$$

where $x_s(\Delta, L)$ and $x_w(\Delta, L)$ are the critical fractions of strong- and weak-bank runners below which perceived fire-sale discount is zero, i.e.

$$x_s(\Delta, L) = \begin{cases} 0 & \text{if } L \leq \frac{\bar{m}}{2} F_\epsilon(\Delta + \underline{\epsilon}), \\ \hat{x}_s : \frac{\bar{m}}{2}\hat{x}_s + \frac{\bar{m}}{2} F_\epsilon\left(\Delta + F_\epsilon^{-1}(\hat{x}_s)\right) = L & \text{if } L > \frac{\bar{m}}{2} F_\epsilon(\Delta + \underline{\epsilon}), \end{cases} \tag{32}$$

$$x_w(\Delta, L) = \begin{cases} \hat{x}_w : \frac{\bar{m}}{2}\hat{x}_w + \frac{\bar{m}}{2} F_\epsilon\left(-\Delta + F_\epsilon^{-1}(\hat{x}_w)\right) = L & \text{if } L < \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon(-\Delta + \bar{\epsilon}), \\ 1 & \text{if } L \geq \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon(-\Delta + \bar{\epsilon}). \end{cases} \tag{33}$$

The run thresholds for strong and weak banks are implicitly given by

$$\int_0^1 \frac{\theta_s^* + \eta}{1 - \bar{m}x} dx - \hat{I}_s(\Delta, l) = 1, \tag{34}$$

$$\int_0^1 \frac{\theta_w^* - \eta}{1 - \bar{m}x} dx - \hat{I}_w(\Delta, l) = 1. \tag{35}$$

As in the baseline model, if cross-bank strategic interactions are nontrivial, run thresholds of strong and weak banks are infinitely close to the average threshold

$$\theta^* = \frac{1}{\int_0^1 \frac{dx}{1-\bar{m}x}} \left[1 + \frac{1}{2}\hat{I}_s(\Delta, L) + \frac{1}{2}\hat{I}_w(\Delta, L)\right], \tag{36}$$

where bank heterogeneity $\Delta = \Delta(L, \eta)$ is implicitly defined by

$$\eta = \frac{1}{2\int_0^1 \frac{dx}{1-\bar{m}x}} \left[\hat{I}_s(\Delta, L) - \hat{I}_w(\Delta, L)\right]. \tag{37}$$

In equilibrium, there exist nontrivial cross-bank interactions if (37) has a solution $\Delta = \Delta(L, \eta) < \bar{\Delta}$.

Below, we prove several lemmas that together imply Proposition 5.

Lemma 2. *Suppose that there exist nontrivial cross-bank strategic interactions and $\Delta = \Delta(L, \eta)$ solves (37). Then it must be that $L < \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon(-\Delta(L, \eta) + \bar{\epsilon})$. Moreover, $\frac{d\Delta}{dL} \geq 0$, with the inequality being strict if $L > 0$ or $\lambda''(\cdot) > 0$.*

Proof. Suppose that L and Δ are such that $L \geq \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon(-\Delta + \bar{\epsilon})$. Then, using the definitions (32) and (33), we get

$$x_s(\Delta, L) = \frac{2L}{\bar{m}} - 1 \geq F_\epsilon(-\Delta + \bar{\epsilon})$$

and $x_w(\Delta, L) = 1$. Then the fire-sale pressure terms (30) and (31) become

$$\hat{I}_s(\Delta, L) = \int_{\min\left\{\frac{2L}{\bar{m}}-1,1\right\}}^{1} \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - L\right)\frac{\bar{m}x}{1-\bar{m}x}dx + \int_0^{\min\left\{\frac{2L}{\bar{m}}-1,1\right\}} \frac{\bar{m}x}{1-\bar{m}x}dx,$$
$$\hat{I}_w(\Delta, L) = \int_0^1 \frac{\bar{m}x}{1-\bar{m}x}dx.$$

In this case, the derivative of the right-hand side of (37) with respect to Δ is 0, that is, the right-hand side of (37) does not depend on Δ. In other words, the fire-sale pressure terms do not depend on bank heterogeneity Δ, i.e. there are no cross-bank strategic uncertainties, which is a contradiction.

Therefore, it must be that $L < \frac{\bar{m}}{2} + \frac{\bar{m}}{2}F_\epsilon(-\Delta + \bar{\epsilon})$. Suppose that $\Delta = \Delta(L, \eta)$ solves (37). Then $L < \frac{\bar{m}}{2} + \frac{\bar{m}}{2}F_\epsilon(-\Delta + \bar{\epsilon})$, which implies $0 \leq x_s(\Delta, L) < F_\epsilon(-\Delta + \bar{\epsilon})$ and $x_s(\Delta, L) \leq x_w(\Delta, L) < 1$. Then, using the expressions for the fire-sale pressure terms (30) and (31), we can straightforwardly establish that

$$\frac{\partial \hat{I}_s}{\partial \Delta} > 0 \quad \text{and} \quad \frac{\partial \hat{I}_w}{\partial \Delta} < 0.$$

Moreover,

$$\frac{\partial \hat{I}_s}{\partial L} - \frac{\partial \hat{I}_w}{\partial L} = -\int_{x_s(\Delta,L)}^{1} \lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - L\right)\frac{\bar{m}x}{1-\bar{m}x}dx +$$
$$\int_{x_w(\Delta,L)}^{1} \lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - L\right)\frac{\bar{m}x}{1-\bar{m}x}dx \leq 0.$$

The latter inequality holds because $x_s(\Delta, L) \leq x_w(\Delta, L)$ and $\lambda(\cdot)$ is weakly convex. Moreover, this inequality is strict if $L > 0$ (because then $x_s(\Delta, L) < x_w(\Delta, L)$) or if $\lambda''(\cdot) > 0$.

By the implicit function theorem, it then follows that

$$\frac{d\Delta}{dL} = -\frac{\frac{\partial \hat{I}_s}{\partial \Delta} - \frac{\partial \hat{I}_w}{\partial \Delta}}{\frac{\partial \hat{I}_s}{\partial L} - \frac{\partial \hat{I}_w}{\partial L}} > 0.$$

□

Lemma 2 implies, in particular, that if Equation (37) has a solution $\Delta(L,\eta)$ for some L, then it also has a solution $\Delta(\tilde{L},\eta) < \Delta(L,\eta)$ for any $\tilde{L} \in [0, L)$.

Recall that $\bar{\eta}$, defined in (23), can be written as

$$\bar{\eta} = \frac{1}{2\int_0^1 \frac{1}{1-\bar{m}x}dx} \int_0^1 \left[\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}\right) - \lambda\left(\frac{\bar{m}}{2}x\right)\right] \frac{\bar{m}x}{1-\bar{m}x}dx.$$

Lemma 3. *For each $\eta \in (0,\bar{\eta})$, there exists $\bar{L} = \bar{L}(\eta)$ such that if $L < \bar{L}$, then cross-bank interactions are nontrivial, $\Delta = \Delta(L,\eta) < \bar{\Delta}$. Moreover, $\bar{L}(\eta)$ is a decreasing function such that $\sup_{\eta\in(0,\bar{\eta})} \bar{L}(\eta) = \bar{m}$ and $\inf_{\eta\in(0,\bar{\eta})} \bar{L}(\eta) = 0$.*

Proof. Lemma 2 implies that if (37) has a solution $\Delta = \Delta(L,\eta)$, then $\Delta = \Delta(L,\eta)$ increases in L. Define the supremum value of $\Delta(L,\eta)$ by $\Delta^*(\eta)$. Two cases are possible.

Case 1: $\Delta^*(\eta) = \bar{\Delta}$.

Suppose that $L = \bar{L}(\eta)$ such that $\Delta = \Delta^*(\eta) = \bar{\Delta}$. By Lemma 2, it must be that

$$\bar{L} \leq \frac{\bar{m}}{2} + \frac{\bar{m}}{2}F_\epsilon\left(-\bar{\Delta} + \bar{\epsilon}\right) = \frac{\bar{m}}{2}. \tag{38}$$

Using the definitions (32) and (33), we find $x_s\left(\bar{\Delta},\bar{L}\right) = 0$ and $x_w\left(\bar{\Delta},\bar{L}\right) = \frac{2\bar{L}}{\bar{m}}$. Plugging these into (37), we obtain

$$\eta = \frac{1}{2\int_0^1 \frac{1}{1-\bar{m}x}dx} \times \left[\int_0^1 \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - \bar{L}\right)\frac{\bar{m}x}{1-\bar{m}x}dx - \int_{\frac{2\bar{L}}{\bar{m}}}^1 \lambda\left(\frac{\bar{m}}{2}x - \bar{L}\right)\frac{\bar{m}x}{1-\bar{m}x}dx - \int_0^{\frac{2\bar{L}}{\bar{m}}} \frac{\bar{m}x}{1-\bar{m}x}dx\right].$$

This equation defines $\bar{L} = \bar{L}(\eta)$ implicitly. It is easy to verify that the right-hand side decreases in $\bar{L}$ if $\lambda(\cdot)$ is weakly convex. Therefore, $\bar{L}(\eta)$ is a decreasing function. Define

$$\hat{\eta} = \frac{1}{2\int_0^1 \frac{1}{1-\bar{m}x}dx} \int_0^1 \left[\lambda\left(\frac{\bar{m}}{2}x\right) - 1\right] \frac{\bar{m}x}{1-\bar{m}x}dx. \tag{39}$$

Note that if $\lambda(\cdot)$ is weakly convex, then $\bar{\eta} > \hat{\eta}$.

It is easy to see that $\bar{L}(\hat{\eta}) = \frac{\bar{m}}{2}$ and $\bar{L}(\bar{\eta}) = 0$. Furthermore, if $\eta < \hat{\eta}$, then $\bar{L}(\eta) > \frac{\bar{m}}{2}$, which contradicts (38).

Case 2: $\Delta^*(\eta) < \bar{\Delta}$.

Suppose that $L = \bar{L}(\eta)$ such that $\Delta = \Delta^*(\eta) < \bar{\Delta}$. By Lemma 2, it must be that

$$\bar{L} \leq \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right). \tag{40}$$

Suppose that this inequality holds strictly. Then, using the definitions (32) and (33), we get $x_s\left(\Delta^*, \bar{L}\right) < F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right)$ and $x_w\left(\Delta^*, \bar{L}\right) < 1$. But then the right-hand side of (37) strictly increases in Δ and strictly decreases in L in the neighborhood of $\left(\Delta^*, \bar{L}\right)$. Therefore, for a given η, a marginal increase in $\bar{L}$ implies a marginal increase in Δ^*. Moreover, such changes do not violate (40). Therefore, Δ^* is not the supremum value of $\Delta(L, \eta)$ for a given η. Therefore, it must be that (40) holds as equality, i.e.,

$$\bar{L} = \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right). \tag{41}$$

Under (41), we get $x_s\left(\Delta^*, \bar{L}\right) = F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right)$ and $x_w\left(\Delta^*, \bar{L}\right) = 1$ from (32) and (33). Plugging these into (37), we obtain

$$\eta = \frac{1}{2\int_0^1 \frac{1}{1-\bar{m}x} dx} \int_{F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right)}^1 \left[\lambda\left(\frac{\bar{m}}{2}\left(x - F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right)\right)\right) - 1\right] \frac{\bar{m}x}{1 - \bar{m}x} dx.$$

This equation defines $\Delta^* = \Delta^*(\eta)$ implicitly. Clearly, $\Delta^*(\eta)$ is an increasing function. From (41) it then follows that $\bar{L}(\eta)$ is a decreasing function. Note that $\Delta^* < \bar{\Delta} = \bar{\epsilon} - \underline{\epsilon} \Leftrightarrow \eta < \hat{\eta}$, where $\hat{\eta}$ is defined by (39). Finally, $\Delta^*(0) = 0$ and $\Delta^*(\hat{\eta}) = \bar{\Delta}$, and so $\bar{L}(0) = \bar{m}$ and $L(\hat{\eta}) = \frac{\bar{m}}{2}$.

The analyses of cases 1 and 2 imply that if $L = \bar{L}(\eta)$ then (37) has a solution $\Delta = \Delta^*(\eta)$. Therefore, by Lemma 2, (37) has a solution $\Delta(L, \eta) < \Delta^*(\eta)$ for any $L \in [0, \bar{L}(\eta))$. $\square$

Finally, we establish

Lemma 4. *Suppose that $\eta \in (0, \bar{\eta})$ and $L < \bar{L}(\eta)$. Then $\frac{\partial \theta^*}{\partial L} < 0$ and $\frac{\partial \theta^*}{\partial \Delta} < 0$.*

Proof. From Lemmas 2 and 3 it follows that if $L < \bar{L}(\eta)$ then $0 \leq x_s(\Delta, L) < F_\epsilon(-\Delta + \bar{\epsilon})$ and $x_s(\Delta, L) \leq x_w(\Delta, L) < 1$. Therefore, using the definitions of the fire-sale pressure terms (30) and (31), we get

$$\frac{\partial \hat{I}_s}{\partial L} < 0 \quad \text{and} \quad \frac{\partial \hat{I}_w}{\partial L} < 0,$$

which together imply that $\frac{\partial \theta^*}{\partial L} < 0$, where θ^* is defined by (36).

In the absence of liquidity injections, $\frac{\partial \theta^*}{\partial \Delta} < 0$ by Proposition 1. Below we show that this result holds if $L \in \left(0, \bar{L}\left(\eta\right)\right)$.

$$\frac{\partial \theta^*}{\partial \Delta} \propto \int_{x_s(\Delta,L)}^{F_\epsilon(-\Delta+\bar{\epsilon})} \lambda' \left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - L\right) f_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) \frac{\bar{m}x}{1-\bar{m}x}dx -$$
$$\int_{\max\{x_w(\Delta,L),F_\epsilon(\Delta+\underline{\epsilon})\}}^{1} \lambda' \left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - L\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) \frac{\bar{m}x}{1-\bar{m}x}dx,$$

where $\propto$ denotes proportionality up to a positive multiplicative term. By changing the variable of integration in the second integral, $y = F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)$, we can rewrite the expression as

$$\frac{\partial \theta^*}{\partial \Delta} \propto \int_{x_s(\Delta,L)}^{F_\epsilon(-\Delta+\bar{\epsilon})} \lambda' \left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - L\right) f_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) \times$$
$$\left(\frac{\bar{m}x}{1-\bar{m}x} - \frac{\bar{m}F_\epsilon(\Delta + F_\epsilon^{-1}(x))}{1-\bar{m}F_\epsilon(\Delta + F_\epsilon^{-1}(x))}\right)dx < 0.$$

□

Proposition 5 follows from Lemmas 2, 3 and 4.

B.5.2 Corollary 3

Proof. The proof of Lemma 3 implies that if $\eta \geq \bar{\eta}$ or $L \geq \bar{L}(\eta)$, then cross-bank interactions are trivial, $\Delta \geq \bar{\Delta}$. Plugging $\Delta \geq \bar{\Delta}$ in (30) and (31), we obtain

$$\hat{I}_s(L) = \int_{x_s(L)}^{1} \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - L\right)\frac{\bar{m}x}{1-\bar{m}x}dx + \int_0^{x_s(L)} \frac{\bar{m}x}{1-\bar{m}x}dx,$$
$$\hat{I}_w(L) = \int_{x_w(L)}^{1} \lambda\left(\frac{\bar{m}}{2}x - L\right)\frac{\bar{m}x}{1-\bar{m}x}dx + \int_0^{x_w(L)} \frac{\bar{m}x}{1-\bar{m}x}dx,$$

where from (32) and (33)

$$x_s(L) = \begin{cases} 0 & \text{if } L \leq \frac{\bar{m}}{2}, \\ \frac{2}{\bar{m}}L - 1 & \text{if } L > \frac{\bar{m}}{2}, \end{cases}$$

$$x_w(L) = \begin{cases} \frac{2}{\bar{m}}L & \text{if } L < \frac{\bar{m}}{2}, \\ 1 & \text{if } L \geq \frac{\bar{m}}{2}. \end{cases}$$

Therefore, the average fragility θ^* no longer depends on Δ,

$$\theta^* = \frac{1}{\int_0^1 \frac{dx}{1-\bar{m}x}} \left(1 + \frac{1}{2}\hat{I}_s(L) + \frac{1}{2}\hat{I}_w(L)\right).$$

Since θ^* no longer depends on Δ, $\frac{\partial\theta^*}{\partial\Delta} = 0$. Nonetheless, liquidity injection reduces fire-sale pressure directly, i.e. $\frac{d\hat{I}_s(L)}{dL} < 0$ and $\frac{d\hat{I}_w(L)}{dL} < 0$. Therefore, liquidity injection reduces fragility only through its direct effect, $\frac{d\theta^*}{dL} = \frac{\partial\theta^*}{\partial L} < 0$. Furthermore, from (34) and (35), $\frac{d\theta_s^*}{dl} = \frac{\partial\theta_s^*}{\partial l} < 0$ and $\frac{d\theta_w^*}{dl} = \frac{\partial\theta_w^*}{\partial l} < 0$. □

B.6 Liquidity buffers

This appendix proves Proposition 6 and Corollary 4. Throughout this appendix, we assume that Assumption 1 holds.

B.6.1 Proposition 6

Define

$$\hat{I}_s(\Delta, l) = \int_{\frac{l}{\bar{m}}}^1 \frac{\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1 - \bar{m}x}dx - \int_0^{\frac{l}{\bar{m}}} \frac{l - \bar{m}x}{1 - \bar{m}x}dx, \tag{42}$$

$$\hat{I}_w(\Delta, l) = \int_{x_w(\Delta,l)}^1 \frac{\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1 - \bar{m}x}dx - \int_0^{x_w(\Delta,l)} \frac{l - \bar{m}x}{1 - \bar{m}x}dx, \tag{43}$$

where $x_w(\Delta, l)$ is the critical fraction of weak-bank runners below which all assets liquidated by weak banks are absorbed by strong banks. It is given by

$$x_w(\Delta, l) = \begin{cases} \hat{x}_w : \frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) = l & \text{if } l < \frac{\bar{m}}{2} + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + \bar{\epsilon}\right), \\ 1 & \text{if } l \geq \frac{\bar{m}}{2} + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + \bar{\epsilon}\right). \end{cases} \tag{44}$$

The run thresholds for strong and weak banks are implicitly given by

$$\int_0^1 \frac{(\theta_s^* + \eta)(1-l)}{1-\bar{m}x} dx - \hat{I}_s(\Delta, l) = 1, \tag{45}$$

$$\int_0^1 \frac{(\theta_w^* - \eta)(1-l)}{1-\bar{m}x} dx - \hat{I}_w(\Delta, l) = 1. \tag{46}$$

As in the baseline model, if cross-bank strategic interactions are nontrivial, run thresholds of strong and weak banks are infinitely close to the average threshold

$$\theta^* = \frac{1}{\int_0^1 \frac{1-l}{1-\bar{m}x} dx} \left(1 + \frac{1}{2}\hat{I}_s(\Delta, l) + \frac{1}{2}\hat{I}_w(\Delta, l)\right), \tag{47}$$

where bank heterogeneity $\Delta = \Delta(l, \eta)$ is implicitly given by

$$\eta = \frac{1}{2\int_0^1 \frac{1-l}{1-\bar{m}x} dx} \left(\hat{I}_s(\Delta, l) - \hat{I}_w(\Delta, l)\right). \tag{48}$$

In equilibrium, there exist nontrivial cross-bank interactions if (48) has a solution $\Delta = \Delta(l, \eta) < \bar{\Delta}$.

Below, we prove several lemmas that together imply Proposition 6.

Lemma 5. *Suppose that there exist nontrivial cross-bank strategic interactions and $\Delta = \Delta(l, \eta)$ solves (48). Then it must be that $l < \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon(-\Delta(l, \eta) + \bar{\epsilon})$. Moreover, $\frac{d\Delta}{dl} \geq 0$, with the inequality being strict if $l > 0$ or $\lambda''(\cdot) > 0$.*

Proof. Suppose that l and Δ are such that

$$l \geq \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon(-\Delta + \bar{\epsilon}). \tag{49}$$

Then Equations (42) and (43) become

$$\hat{I}_s(\Delta, l) = \int_{\frac{l}{\bar{m}}}^1 \frac{\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x} dx - \int_0^{\frac{l}{\bar{m}}} \frac{l - \bar{m}x}{1-\bar{m}x} dx, \tag{50}$$

$$\hat{I}_w(\Delta, l) = -\int_0^1 \frac{l-\bar{m}x}{1-\bar{m}x} dx.$$

Note that (49) implies that $\frac{l}{\bar{m}} \geq F_\epsilon(-\Delta + \bar{\epsilon})$. Therefore, if $x \geq \frac{l}{\bar{m}}$ then $F_\epsilon(\Delta + F_\epsilon^{-1}(x)) = 1$, and so the right-hand side of (50) does not depend on Δ. Hence, the right-hand side of (48) also does not depend on Δ, i.e. there are no cross-bank strategic uncertainties, which is a contradiction.

Suppose now that l and Δ are such that $l < \frac{\bar{m}}{2}+\frac{\bar{m}}{2}F_{\epsilon}\left(-\Delta+\bar{\epsilon}\right)$. Then $\frac{l}{\bar{m}} \leq x_w\left(\Delta,l\right) < 1$ and $\hat{I}_w\left(\Delta,l\right)$ is a decreasing function of Δ. If $\frac{l}{\bar{m}} \geq F_{\epsilon}\left(-\Delta+\bar{\epsilon}\right)$ then $\hat{I}_s\left(\Delta,l\right)$ does not depend on Δ. If $\frac{l}{\bar{m}} < F_{\epsilon}\left(-\Delta+\bar{\epsilon}\right)$ then $\hat{I}_s\left(\Delta,l\right)$ is an increasing function of Δ. Denote the right-hand side of (48) by $T\left(\Delta,l\right)$. We have $\frac{\partial T}{\partial\Delta} > 0$. By the implicit function theorem,

$$\frac{d\Delta}{dl} = -\frac{\frac{\partial T}{\partial l}}{\frac{\partial T}{\partial\Delta}} = \frac{1}{\frac{\partial T}{\partial\Delta}}\frac{T_1+T_2}{2\left(1-l\right)^2\int_0^1\frac{dx}{1-\bar{m}x}},$$

where

$$T_1 = \left(1-l\right)\times\left(\int_{\frac{l}{\bar{m}}}^{x_w(\Delta,l)}\frac{\lambda'\left(\frac{\bar{m}}{2}x+\frac{\bar{m}}{2}F_{\epsilon}\left(\Delta+F_{\epsilon}^{-1}\left(x\right)\right)-l\right)\left(\bar{m}x-l\right)}{1-\bar{m}x}dx+\right.$$
$$\left.\int_{x_w(\Delta,l)}^{1}\frac{\left(\lambda'\left(\frac{\bar{m}}{2}x+\frac{\bar{m}}{2}F_{\epsilon}\left(\Delta+F_{\epsilon}^{-1}\left(x\right)\right)-l\right)-\lambda'\left(\frac{\bar{m}}{2}x+\frac{\bar{m}}{2}F_{\epsilon}\left(-\Delta+F_{\epsilon}^{-1}\left(x\right)\right)-l\right)\right)\left(\bar{m}x-l\right)}{1-\bar{m}x}dx\right),$$

and

$$T_2 = \int_{\frac{l}{\bar{m}}}^{x_w(\Delta,l)}\left[\lambda\left(\frac{\bar{m}}{2}x+\frac{\bar{m}}{2}F_{\epsilon}\left(\Delta+F_{\epsilon}^{-1}\left(x\right)\right)-l\right)-1\right]dx+$$
$$\int_{x_w(\Delta,l)}^{1}\left(\lambda\left(\frac{\bar{m}}{2}x+\frac{\bar{m}}{2}F_{\epsilon}\left(\Delta+F_{\epsilon}^{-1}\left(x\right)\right)-l\right)-\lambda\left(\frac{\bar{m}}{2}x+\frac{\bar{m}}{2}F_{\epsilon}\left(-\Delta+F_{\epsilon}^{-1}\left(x\right)\right)-l\right)\right)dx.$$

Note that $T_2 \geq 0$ and $T_1 \geq 0$ under Assumption 1. Therefore, $\frac{d\Delta}{dl} \geq 0$. Furthermore, this inequality is strict if $l > 0$ (because then $x_w\left(\Delta,l\right) > \frac{l}{\bar{m}}$ and $T_2 > 0$) or if $\lambda''(\cdot) > 0$. □

Lemma 5 implies, in particular, that if Equation (48) has a solution $\Delta\left(l,\eta\right)$ for some $l > 0$, then it also has a solution $\Delta(\tilde{l},\eta) < \Delta\left(l,\eta\right)$ for any $\tilde{l} \in [0,l)$.

Recall that $\bar{\eta}$, defined in (23), can be written as

$$\bar{\eta} = \frac{1}{2\int_0^1\frac{1}{1-\bar{m}x}dx}\int_0^1\left[\lambda\left(\frac{\bar{m}}{2}x+\frac{\bar{m}}{2}\right)-\lambda\left(\frac{\bar{m}}{2}x\right)\right]\frac{\bar{m}x}{1-\bar{m}x}dx.$$

Lemma 6. *For each $\eta \in (0,\bar{\eta})$, there exists $\bar{l} = \bar{l}\left(\eta\right)$ such that if $l < \bar{l}$, then cross-bank interactions are nontrivial, $\Delta = \Delta\left(l,\eta\right) < \bar{\Delta}$. Moreover, $\bar{l}\left(\eta\right)$ is a decreasing function such that $\sup_{\eta\in(0,\bar{\eta})}\bar{l}\left(\eta\right) = \bar{m}$ and $\inf_{\eta\in(0,\bar{\eta})}\bar{l}\left(\eta\right) = 0$.*

Proof. Lemma 5 implies that if (48) has a solution $\Delta = \Delta(l, \eta)$, then $\Delta(l, \eta)$ increases in l. Define supremum value of $\Delta(l, \eta)$ by $\Delta^*(\eta)$. Two cases are possible.

Case 1: $\Delta^*(\eta) = \bar{\Delta}$.

Suppose that $l = \bar{l}(\eta)$ such that $\Delta = \Delta^*(\eta) = \bar{\Delta}$. By Lemma 5, it must be that

$$\bar{l} \leq \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon\left(-\bar{\Delta} + \bar{\epsilon}\right) = \frac{\bar{m}}{2}. \tag{51}$$

Using the definition (44), we find $x_w\left(\bar{\Delta}, \bar{l}\right) = \frac{2\bar{l}}{\bar{m}}$. Then (48) can be rewritten as

$$\eta = \frac{\int_{\frac{l}{\bar{m}}}^{1} \frac{\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - \bar{l}\right)\left(\bar{m}x - \bar{l}\right)}{1-\bar{m}x} dx - \int_{\frac{2\bar{l}}{\bar{m}}}^{1} \frac{\lambda\left(\frac{\bar{m}}{2}x - \bar{l}\right)\left(\bar{m}x - \bar{l}\right)}{1-\bar{m}x} dx + \int_{\frac{l}{\bar{m}}}^{\frac{2\bar{l}}{\bar{m}}} \frac{\bar{l} - \bar{m}x}{1-\bar{m}x} dx}{2\int_0^1 \frac{1-\bar{l}}{1-\bar{m}x} dx}.$$

This equation defines $\bar{l} = \bar{l}(\eta)$ implicitly. By the implicit function theorem,

$$\frac{d\bar{l}}{d\eta} = -\frac{1}{2\left(1-\bar{l}\right)\int_0^1 \frac{dx}{1-\bar{m}x}} \times$$
$$\left[\frac{1}{1-\bar{l}}\left(\int_{\frac{l}{\bar{m}}}^{\frac{2\bar{l}}{\bar{m}}} \left(\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - \bar{l}\right) - 1\right) dx + \int_{\frac{2\bar{l}}{\bar{m}}}^{1} \left(\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - \bar{l}\right) - \lambda\left(\frac{\bar{m}}{2}x - \bar{l}\right)\right) dx\right) + \right.$$
$$\left. \int_{\frac{l}{\bar{m}}}^{\frac{2\bar{l}}{\bar{m}}} \frac{\lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - \bar{l}\right)\left(\bar{m}x - \bar{l}\right)}{1-\bar{m}x} dx + \int_{\frac{2\bar{l}}{\bar{m}}}^{1} \frac{\left(\lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - \bar{l}\right) - \lambda'\left(\frac{\bar{m}}{2}x - \bar{l}\right)\right)\left(\bar{m}x - \bar{l}\right)}{1-\bar{m}x} dx \right] < 0.$$

Therefore, $\bar{l}(\eta)$ is a decreasing function. Define

$$\hat{\eta} = \frac{1}{2\int_0^1 \frac{1-\frac{\bar{m}}{2}}{1-\bar{m}x} dx} \left(\int_{\frac{1}{2}}^{1} \frac{\lambda\left(\frac{\bar{m}}{2}x\right)\left(\bar{m}x - \frac{\bar{m}}{2}\right)}{1-\bar{m}x} dx + \int_{\frac{1}{2}}^{1} \frac{\frac{\bar{m}}{2} - \bar{m}x}{1-\bar{m}x} dx\right). \tag{52}$$

It is easy to see that $\bar{l}(\hat{\eta}) = \frac{\bar{m}}{2}$ and $\bar{l}(\bar{\eta}) = 0$. Furthermore, if $\eta < \hat{\eta}$, then $\bar{l}(\eta) > \frac{\bar{m}}{2}$, which contradicts (51).

Case 2: $\Delta^*(\eta) < \bar{\Delta}$.

Suppose that $l = \bar{l}(\eta)$ such that $\Delta = \Delta^*(\eta) < \bar{\Delta}$. By Lemma 5, it must be that

$$\bar{l} \leq \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right). \tag{53}$$

Suppose that this inequality holds strictly. Then the definition (44) implies that $x_w\left(\Delta^*, \bar{l}\right) <$

1. But then the right-hand side of (48) strictly increases in Δ and strictly decreases in l in the neighborhood of $\left(\Delta^*, \bar{l}\right)$. Therefore, for a given η, a marginal increase in $\bar{l}$ implies a marginal increase in Δ^*. Moreover, such changes do not violate (53). Therefore, Δ^* is not the supremum value of $\Delta(l, \eta)$ for a given η, and so (53) must hold as equality, i.e.,

$$\bar{l} = \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right) > \frac{\bar{m}}{2}. \tag{54}$$

Under (54), $x_w\left(\Delta^*, \bar{l}\right) = 1$ and $\frac{\bar{l}}{\bar{m}} = \frac{1}{2} + \frac{1}{2} F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right) > F_\epsilon\left(-\Delta^* + \bar{\epsilon}\right)$. Then (48) can be rewritten as

$$\eta = \frac{1}{2\int_0^1 \frac{1-\bar{l}}{1-\bar{m}x} dx} \int_{\frac{\bar{l}}{\bar{m}}}^1 \left[\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - \bar{l}\right) - 1\right] \frac{\bar{m}x - \bar{l}}{1 - \bar{m}x} dx.$$

This equation defines $\bar{l} = \bar{l}(\eta)$ implicitly. By the implicit function theorem,

$$\frac{d\bar{l}}{d\eta} = -\frac{\left(1-\bar{l}\right)\int_{\frac{\bar{l}}{\bar{m}}}^1 \lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - \bar{l}\right) \frac{\bar{m}x - \bar{l}}{1-\bar{m}x} dx + \int_{\frac{\bar{l}}{\bar{m}}}^1 \left[\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - \bar{l}\right) - 1\right] dx}{2\left(1-\bar{l}\right)^2 \int_0^1 \frac{dx}{1-\bar{m}x}} < 0.$$

Therefore, $\bar{l}(\eta)$ is a decreasing function. Clearly, $\bar{l}(0) = \bar{m}$ and $\bar{l}(\hat{\eta}) = \frac{\bar{m}}{2}$, where $\hat{\eta}$ is given by (52). Furthermore, if $\eta > \hat{\eta}$ then $\bar{l}(\eta) < \frac{\bar{m}}{2}$, which contradicts (54).

The analyses of cases 1 and 2 imply that if $l = \bar{l}(\eta)$ then (48) has a solution $\Delta = \Delta^*(\eta)$. Therefore, by Lemma 5, (48) has a solution $\Delta(l, \eta) < \Delta^*(\eta)$ for any $l \in [0, \bar{l}(\eta))$. □

Finally, we establish

Lemma 7. *Suppose that $\eta \in (0, \bar{\eta})$ and $l < \bar{l}(\eta)$. Then $\frac{\partial \theta^*}{\partial l} < 0$ and $\frac{\partial \theta^*}{\partial \Delta} < 0$.*

Proof. From Lemmas 5 and 6 it follows that if $l < \bar{l}(\eta)$ then cross-bank strategic interactions are nontrivial and $l < \frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon\left(-\Delta + \bar{\epsilon}\right)$. By definition (44), $\frac{l}{\bar{m}} \le x_w(\Delta, l) < 1$.

Using (42) and (43), we can write the derivative of the average threshold (47) with respect to l as

$$\frac{\partial\theta^*}{\partial l} = -\frac{1}{2(1-l)\int_0^1 \frac{1}{1-\bar{m}x}dx} \times \left[\int_{\frac{l}{\bar{m}}}^1 \frac{\lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x}dx + \right.$$
$$\int_{x_w(\Delta,l)}^1 \frac{\lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x}dx +$$
$$\frac{1}{1-l}\int_{\frac{l}{\bar{m}}}^1 \left(\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - l\right) - 1\right)dx +$$
$$\left.\frac{1}{1-l}\int_{x_w(\Delta,l)}^1 \left(\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - l\right) - 1\right)dx\right] < 0.$$

Next, suppose that $\frac{l}{\bar{m}} \geq F_\epsilon(-\Delta + \bar{\epsilon})$. Then the right-hand side of (42) does not depend on Δ, and so

$$\frac{\partial\theta^*}{\partial\Delta} = -\frac{\int_{x_w(\Delta,l)}^1 \frac{\frac{\bar{m}}{2}f_\epsilon\left(-\Delta+F_\epsilon^{-1}(x)\right)\lambda'\left(\frac{\bar{m}}{2}x+\frac{\bar{m}}{2}F_\epsilon\left(-\Delta+F_\epsilon^{-1}(x)\right)-l\right)(\bar{m}x-l)}{1-\bar{m}x}dx}{2\int_0^1 \frac{1-l}{1-\bar{m}x}dx} < 0.$$

Suppose now that $\frac{l}{\bar{m}} < F_\epsilon(-\Delta + \bar{\epsilon})$. From (44) it then follows that $x_w(\Delta, l) \leq F_\epsilon\left(\Delta + F_\epsilon^{-1}\left(\frac{l}{\bar{m}}\right)\right) < 1$. Then

$$\frac{\partial\theta^*}{\partial\Delta} \propto \int_{\frac{l}{\bar{m}}}^1 \frac{\frac{\bar{m}}{2}f_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right)\lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x}dx -$$
$$\int_{x_w(\Delta,l)}^1 \frac{\frac{\bar{m}}{2}f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x}dx,$$

where $\propto$ denotes proportionality up to a positive multiplicative term. By changing the variable of integration in the first integral, $y = F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right)$, we can rewrite the expression as

$$\frac{\partial\theta^*}{\partial\Delta} \propto \int_{F_\epsilon\left(\Delta+F_\epsilon^{-1}\left(\frac{l}{\bar{m}}\right)\right)}^1 \frac{\bar{m}}{2}f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\lambda'\left(\frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) + \frac{\bar{m}}{2}x - l\right) \times$$
$$\left[\frac{\bar{m}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - l}{1 - \bar{m}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)} - \frac{\bar{m}x - l}{1-\bar{m}x}\right]dx -$$
$$\int_{x_w(\Delta,L)}^{F_\epsilon\left(\Delta+F_\epsilon^{-1}\left(\frac{l}{\bar{m}}\right)\right)} \frac{\frac{\bar{m}}{2}f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x}dx < 0.$$

□

Proposition 6 follows from Lemmas 5, 6 and 7.

B.6.2 Corollary 4

Proof. The proof of Lemma 6 implies that if $\eta \geq \bar{\eta}$ or $l \geq \bar{l}(\eta)$, then cross-bank interactions are trivial, $\Delta \geq \bar{\Delta}$. Plugging $\Delta \geq \bar{\Delta}$ in (42) and (43), we obtain

$$\hat{I}_s(l) = \int_{\frac{l}{\bar{m}}}^{1} \frac{\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2} - l\right)(\bar{m}x - l)}{1 - \bar{m}x} dx - \int_0^{\frac{l}{\bar{m}}} \frac{l - \bar{m}x}{1 - \bar{m}x} dx,$$
$$\hat{I}_w(l) = \int_{x_w(l)}^{1} \frac{\lambda\left(\frac{\bar{m}}{2}x - l\right)(\bar{m}x - l)}{1 - \bar{m}x} dx - \int_0^{x_w(l)} \frac{l - \bar{m}x}{1 - \bar{m}x} dx,$$

where from (44)

$$x_w(l) = \begin{cases} \frac{2l}{\bar{m}} & \text{if } l < \frac{\bar{m}}{2}, \\ 1 & \text{if } l \geq \frac{\bar{m}}{2}. \end{cases}$$

Therefore, the average fragility θ^* no longer depends on heterogeneity Δ,

$$\theta^* = \frac{1}{\int_0^1 \frac{1-l}{1-\bar{m}x} dx} \left(1 + \frac{1}{2}\hat{I}_s(l) + \frac{1}{2}\hat{I}_w(l)\right).$$

Since θ^* no longer depends on Δ, $\frac{\partial\theta^*}{\partial\Delta} = 0$. Nonetheless, liquidity buffer requirement reduces fire-sale pressure directly, i.e. $\frac{d\hat{I}_s(l)}{dl} < 0$ and $\frac{d\hat{I}_w(l)}{dl} < 0$. Therefore, using the same approach as in proof of Lemma 7, one can show that the liquidity buffer requirement reduces fragility only through its direct effect, $\frac{d\theta^*}{dl} = \frac{\partial\theta^*}{\partial l} < 0$. Furthermore, from (45) and (46), $\frac{d\theta^*_s}{dl} = \frac{\partial\theta^*_s}{\partial l} < 0$ and $\frac{d\theta^*_w}{dl} = \frac{\partial\theta^*_w}{\partial l} < 0$. □

C Robustness and model extensions

C.1 Inefficient asset management

In Section 2.2, the fire-sale discount is a result of liquidity shortage in the asset market. In this section, we consider an alternative setup of the asset market to illustrate that the fire-sale discount can arise when liquidity is abundant but outside investors are less efficient in managing assets than banks.

Assume that outside investors have abundant liquidity, i.e., $g(L) = L$. However, they are less efficient in managing assets. In particular, under banks' management, in the absence of premature liquidations, a portfolio $\{k_i\}_{i\in[0,1]}$ generates $y \equiv \int z_i k_i di$ at $t = 2$. In contrast, if the same portfolio is managed by outside investors, the return is subject to a discount: instead of receiving y, outside investors only get $f(y)$, where $f(y) < y$ for all $y > 0$ and $f(0) = 0$. In addition, we assume that $f'(\cdot) > 0$ and $f''(\cdot) < 0$ so that outsiders' inefficiency in production increases in the amount of assets they absorb. Furthermore, we assume that $yf'(y)$ is increasing in y to guarantee equilibrium uniqueness in the asset market at $t = 1$. These assumptions on $f(\cdot)$ are typical in the literature on fire sales (e.g., Lorenzoni, 2008).

The outside investors' problem therefore becomes

$$\max_{\{k_i\}_{i\in[0,1]}} f\left(\int z_i k_i di\right) - \int p_i k_i di.$$

The first-order conditions of the outside investors' problem imply that

$$p_i = \frac{\partial f(y)}{\partial y} z_i \quad \forall i \in [0,1], \tag{55}$$

where $y \equiv \int z_i k_i di$. After imposing the market clearing conditions, $k_i = \frac{m_i}{p_i}$ $\forall i \in [0,1]$, we obtain

$$m_i = z_i k_i \frac{\partial f(y)}{\partial y} \Rightarrow m \equiv \int m_i di = y\frac{\partial f(y)}{\partial y}.$$

Since by assumption $yf'(y)$ is an increasing function of y, there is a unique solution $y = h(m)$ to the equation above. Moreover, $h'(\cdot) > 0$. Plugging this into (55), we obtain the equilibrium prices of the same form as in Lemma 1,

$$p_i(z_i, m) = z_i \frac{m}{h(m)} = \frac{z_i}{\lambda(m)},$$

where $\lambda(m) \equiv \frac{h(m)}{m}$.

Moreover, the liquidation price for any asset i is a decreasing function of the total mass of early withdrawers m. Indeed, using (55), we can write

$$p_i(z_i, m) = z_i \frac{\partial f(y)}{\partial y}\bigg|_{y=h(m)} \Rightarrow \frac{\partial p_i}{\partial m} = z_i \frac{\partial^2 f(y)}{\partial y^2}\bigg|_{y=h(m)} \times \frac{\partial h(m)}{\partial m} < 0.$$

C.2 General payoff function

Consider a general incremental payoff function $\pi(z_i, m_i, m)$ of an investor of bank i that chooses not to withdraw her funds early. It depends on her bank's productivity z_i, mass of runners on her bank m_i, and overall mass of runners in the whole economy m. In the main model, $\pi(z_i, m_i, m) = \frac{z_i - \lambda(m) m_i}{1-m_i} - 1$. In Section 3.3, we consider $\pi(z_i, m_i, m) = z_i \pi_1(m_i) + \pi_2(m_i, m)$ to illustrate the role of two complementarities. In this appendix, we do not assume a specific functional form for $\pi(\cdot,\cdot,\cdot)$. We denote partial derivatives of the $\pi(\cdot,\cdot,\cdot)$ function by subscripts. We assume that the payoff function is smooth and that it satisfies the following monotonicity properties: $\pi_z \equiv \frac{\partial \pi}{\partial z_i} > 0$, $\pi_m \equiv \frac{\partial \pi}{\partial m} < 0$.

Below, we show that if $\pi_{m_i m} \leq 0, \pi_{zm} \geq 0, \pi_{zz} \geq 0$, our main result in Proposition 1 holds, i.e. $\frac{\partial \theta^*}{\partial \Delta} \leq 0$. We focus on the case of nontrivial cross-bank strategic interactions, that is, $\Delta \equiv \lim\limits_{\sigma \to 0} \frac{\theta_w^* - \theta_s^*}{\sigma} \in \left(0, \bar{\Delta}\right)$. The analogues of Equations (7) and (8) from the main text are

$$\int_0^1 \pi\left(\theta_s^* + \eta, \bar{m}x, \frac{\bar{m}}{2}x + \frac{\bar{m}}{2} F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right)\right) dx = 0,$$
$$\int_0^1 \pi\left(\theta_w^* - \eta, \bar{m}x, \frac{\bar{m}}{2}x + \frac{\bar{m}}{2} F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\right) dx = 0.$$

Define

$$m_{tot}(x, \Delta) = \frac{\bar{m}}{2}x + \frac{\bar{m}}{2} F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right).$$

Note that $m_{tot}(x, \Delta) > \bar{m}x > m_{tot}(x, -\Delta)$. By the implicit function theorem,

$$\frac{\partial \theta_s^*}{\partial \Delta} = -\frac{\frac{\bar{m}}{2} \int_0^{F_\epsilon(\bar{\epsilon}-\Delta)} \pi_m\left(\theta_s^* + \eta, \bar{m}x, m_{tot}(x, \Delta)\right) f_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) dx}{\int_0^1 \pi_z\left(\theta_s^* + \eta, \bar{m}x, m_{tot}(x, \Delta)\right) dx} > 0,$$
$$\frac{\partial \theta_w^*}{\partial \Delta} = -\frac{-\frac{\bar{m}}{2} \int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \pi_m\left(\theta_w^* - \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) dx}{\int_0^1 \pi_z\left(\theta_w^* - \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) dx} < 0.$$

Then the impact of Δ on the average threshold $\theta^* \equiv \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$ is

$$\frac{\partial \theta^*}{\partial \Delta} = \frac{1}{2}\frac{\partial \theta_s^*}{\partial \Delta} + \frac{1}{2}\frac{\partial \theta_w^*}{\partial \Delta} = \frac{1}{2}\frac{\partial \theta_s^*}{\partial \Delta}\left(\frac{\frac{\partial \theta_w^*}{\partial \Delta}}{\frac{\partial \theta_s^*}{\partial \Delta}} + 1\right).$$

When $\sigma \to 0$ and $\Delta < \bar{\Delta}$, $\theta_s^* \to \theta^*$ and $\theta_w^* \to \theta^*$, so that

$$\frac{\partial \theta^*}{\partial \Delta} \propto \frac{\frac{\partial \theta_w^*}{\partial \Delta}\Big|_{\theta_w^*=\theta^*}}{\frac{\partial \theta_s^*}{\partial \Delta}\Big|_{\theta_s^*=\theta^*}} + 1 = A_1 \times A_2 + 1,$$

where $\propto$ denotes proportionality up to a positive multiplicative term,

$$\begin{aligned} A_1 = & \frac{\int_0^1 \pi_z\left(\theta^* + \eta, \bar{m}x, m_{tot}(x, \Delta)\right) dx}{\int_0^1 \pi_z\left(\theta^* - \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) dx} \overset{\pi_{zm}\geq 0}{\geq} \\ & \frac{\int_0^1 \pi_z\left(\theta^* + \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) dx}{\int_0^1 \pi_z\left(\theta^* - \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) dx} \overset{\pi_{zz}\geq 0}{\geq} 1, \end{aligned}$$

and

$$\begin{aligned} A_2 = & -\frac{\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \pi_m\left(\theta^* - \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) dx}{\int_0^{F_\epsilon(\bar{\epsilon}-\Delta)} \pi_m\left(\theta^* + \eta, \bar{m}x, m_{tot}(x, \Delta)\right) f_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) dx} = \\ & -\frac{\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \pi_m\left(\theta^* - \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) dx}{\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \pi_m\left(\theta^* + \eta, \bar{m}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right), m_{tot}(x, -\Delta)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) dx} \overset{\pi_{zm}\geq 0}{\leq} \\ & -\frac{\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \pi_m\left(\theta^* + \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) dx}{\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \pi_m\left(\theta^* + \eta, \bar{m}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right), m_{tot}(x, -\Delta)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) dx} \overset{\pi_{m_im}\leq 0}{\leq} \\ & -\frac{\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \pi_m\left(\theta^* + \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) dx}{\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \pi_m\left(\theta^* + \eta, \bar{m}x, m_{tot}(x, -\Delta)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) dx} = -1. \end{aligned}$$

Therefore, $\frac{\partial \theta^*}{\partial \Delta} \leq 0$. Moreover, if one of the inequalities $\{\pi_{m_im} \leq 0, \pi_{zm} \geq 0, \pi_{zz} \geq 0\}$ holds strictly, then $\frac{\partial \theta^*}{\partial \Delta} < 0$.

In the micro-founded case considered in the main text $\pi_{zz} = 0$, $\pi_{zm} = 0$ and $\pi_{m_im} < 0$. Therefore, the crucial underlying economic mechanism behind Propositions 1 and 2 is mutually reinforcing within- and cross-bank complementarities.

C.3 Finite signal precision

In this appendix, we allow the precision of investors' private signals about the aggregate fundamental to be finite, that is, we allow the signal noise σ to be not infinitely close to zero. To keep the analysis tractable, we maintain the assumption that signals are infinitely more precise than any prior information that investors have about the aggregate fundamental θ. Formally, this implies an uninformative prior about θ. In Appendix C.3.2, we consider a numerical example to demonstrate that the results hold even if the prior is informative.[27]

When the prior is uninformative, there exists a unique threshold equilibrium in which investors of strong and weak banks withdraw prematurely if and only if their signals are below θ_s^* and θ_w^*, respectively. With a non-negligible σ, the equations defining the run thresholds (7) and (8) become

$$\begin{aligned}\theta_s^* + \eta =& \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + I_s(\Delta) + \sigma\int_0^1 \frac{F_\epsilon^{-1}(x)}{1-\bar{m}x}dx\right),\\ \theta_w^* - \eta =& \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + I_w(\Delta) + \sigma\int_0^1 \frac{F_\epsilon^{-1}(x)}{1-\bar{m}x}dx\right),\end{aligned}$$

where the term $\sigma\int_0^1 \frac{F_\epsilon^{-1}(x)}{1-\bar{m}x}dx$ arises because signals are not infinitely close to the aggregate fundamental when σ is non-negligible; the fire-sale pressure terms $I_s(\Delta)$ and $I_w(\Delta)$ are defined by (9) and (10), respectively; and $\Delta \equiv \frac{\theta_w^*-\theta_s^*}{\sigma}$ is bank heterogeneity.

Recall that in the main model with negligible σ, weak and strong banks' run thresholds are infinitely close to the average run threshold when bank heterogeneity is not too large ($\Delta < \bar{\Delta}$) and cross-bank strategic interactions are nontrivial. In contrast, if σ is non-negligible, the run thresholds θ_s^* and θ_w^* are different as long as bank assets are not entirely the same, i.e. $\eta \neq 0$. In what follows, we first characterize how the average fragility of the financial system, $\theta^* = \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$, depends on Δ and then consider fragilities of strong and weak banks separately.

The average fragility of the financial system is

$$\theta^* \equiv \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^* = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + \frac{1}{2}I_s(\Delta) + \frac{1}{2}I_w(\Delta) + \sigma\int_0^1 \frac{F_\epsilon^{-1}(x)}{1-\bar{m}x}dx\right), \quad (56)$$

where Δ is implicitly defined by

$$\eta = \frac{\sigma\Delta}{2} + \frac{1}{2\int_0^1 \frac{dx}{1-\bar{m}x}}\left(I_s(\Delta) - I_w(\Delta)\right).$$

These two equations generalize (14) and (15) to the case of finite signal precision. Analogous to our benchmark analysis, average fragility declines in bank heterogeneity as long

[27] As discussed in detail by Morris and Shin (2003), for equilibrium to be unique in global games settings, private signals should be sufficiently more precise than the prior information.

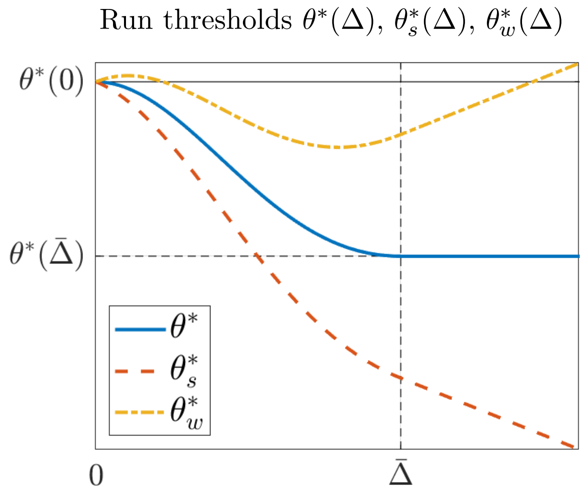

Figure 4: Run thresholds of strong (θ_s^*) and weak (θ_w^*) banks, and their average ($\theta^* = \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$) as functions of bank heterogeneity Δ. Parametrization: $\bar{m} = 0.55$, $\lambda(m) = 1 + m^2$, $F_\epsilon(\cdot)$ is truncated standard normal over $[-1, 1]$, $\sigma = 0.0075$, η varies from 0 to 0.05.

as there are nontrivial cross-bank strategic interactions.

Proposition 7. *In the model with finitely precise signals about the aggregate fundamental θ, the average run threshold $\theta^*(\Delta)$ is a decreasing function of Δ when $\Delta < \bar{\Delta}$. When $\Delta \geq \bar{\Delta}$, $\theta^*(\Delta)$ is constant.*

Proof. See Appendix C.3.1. □

Figure 4 illustrates Proposition 7. When cross-bank interactions are nontrivial, $\Delta < \bar{\Delta}$, the average run threshold (the solid blue line) declines in heterogeneity. In the region where cross-bank strategic uncertainty disappears, $\Delta \geq \bar{\Delta}$, the threshold does not depend on heterogeneity.

The run thresholds of strong and weak banks can be written as

$$\theta_s^* = \theta^*(\Delta) - \frac{1}{2}\sigma\Delta,$$
$$\theta_w^* = \theta^*(\Delta) + \frac{1}{2}\sigma\Delta.$$

When signals have finite precision, individual and average thresholds are no longer infinitely close. Clearly, strong banks become less fragile as Δ increases. However, this does not always hold for weak banks. Nonetheless, as Δ increases, the force that alleviates fire-sale pressure on weak banks and stabilizes them is still at work. If σ is finite but sufficiently small, an increase in bank heterogeneity stabilizes all banks, including the weak ones. The reason is that, unless σ is too large, θ_w^* stays close to θ^* and thus tends to decline with bank heterogeneity. To illustrate that, Figure 4 shows the run thresholds θ_s^* and θ_w^* as functions of bank heterogeneity. As is clear from the graph, if bank heterogeneity is neither too small or too large, an increase in it drives down both run thresholds. Therefore, our key result that bank heterogeneity undermines stability carries through.

C.3.1 Proof of Proposition 7

Proof. The average fragility is given by Equation (56):

$$\theta^* = \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^* = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + \frac{1}{2}I_s(\Delta) + \frac{1}{2}I_w(\Delta) + \sigma\int_0^1 \frac{F_\epsilon^{-1}(x)}{1-\bar{m}x}dx\right).$$

Therefore,

$$\frac{\partial\theta^*}{\partial\Delta} \propto \frac{\partial\left(I_s(\Delta)+I_w(\Delta)\right)}{\partial\Delta},$$

where $\propto$ denotes proportionality up to a positive multiplicative term.

From Equations (9) and (10) defining $I_s(\Delta)$ and $I_w(\Delta)$, it is clear that these terms do not depend on Δ when $\Delta \geq \bar{\Delta}$. If $\Delta < \bar{\Delta}$, we have

$$\begin{aligned}
&\frac{\partial\left(I_s(\Delta)+I_w(\Delta)\right)}{\partial\Delta} \propto \\
&\int_0^{F_\epsilon(\bar{\epsilon}-\Delta)} \lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right)\right) f_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right)\frac{\bar{m}x}{1-\bar{m}x}dx - \\
&\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\frac{\bar{m}x}{1-\bar{m}x}dx = \\
&\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \lambda'\left(\frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) + \frac{\bar{m}}{2}x\right)\frac{\bar{m}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)}{1-\bar{m}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)} f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)dx - \\
&\int_{F_\epsilon(\underline{\epsilon}+\Delta)}^1 \lambda'\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\right) f_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)\frac{\bar{m}x}{1-\bar{m}x}dx < 0,
\end{aligned}$$

where we change the variable of integration $x \to F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right)$ in the first integral.

□

C.3.2 Informative prior

In this appendix, we explore the case in which standard deviation of signal noise σ is finite and the prior about the aggregate fundamental θ is informative. Morris and Shin (2003) show that the results obtained under the improper prior assumption can be continuously extended to the case in which signals are sufficiently, but not necessarily infinitely, more precise than the prior. In what follows, we assume that signals are sufficiently precise, such that the model features unique threshold equilibrium. In particular, the run thresholds of strong- and weak-bank investors solve

$$\frac{\int \pi\left(\theta+\eta, \bar{m}F_\epsilon\left(\frac{\theta_s^*-\theta}{\sigma}\right), \frac{1}{2}\bar{m}F_\epsilon\left(\frac{\theta_s^*-\theta}{\sigma}\right)+\frac{1}{2}\bar{m}F_\epsilon\left(\frac{\theta_w^*-\theta}{\sigma}\right)\right) f_\theta\left(\theta\right) f_\epsilon\left(\frac{\theta_s^*-\theta}{\sigma}\right) d\theta}{\int f_\theta\left(\theta\right) f_\epsilon\left(\frac{\theta_s^*-\theta}{\sigma}\right) d\theta} = 0,$$

$$\frac{\int \pi\left(\theta-\eta, \bar{m}F_\epsilon\left(\frac{\theta_w^*-\theta}{\sigma}\right), \frac{1}{2}\bar{m}F_\epsilon\left(\frac{\theta_s^*-\theta}{\sigma}\right)+\frac{1}{2}\bar{m}F_\epsilon\left(\frac{\theta_w^*-\theta}{\sigma}\right)\right) f_\theta\left(\theta\right) f_\epsilon\left(\frac{\theta_w^*-\theta}{\sigma}\right) d\theta}{\int f_\theta\left(\theta\right) f_\epsilon\left(\frac{\theta_w^*-\theta}{\sigma}\right) d\theta} = 0,$$

where investors of weak and strong banks run if their signals are below θ_w^* and θ_s^*, respectively; $\pi\left(z_i, m_i, m\right)$ is the net benefit of not running defined in (2); $f_\theta(\cdot)$ and $f_\epsilon(\cdot)$ are probability density functions of the prior and signal noise, respectively.

We consider the following parametrization: $\bar{m} = 0.55$; $\lambda(m) = 1 + 2m^2$; noise distribution is standard normal; $\sigma = 0.005$; the prior is normal with a mean of 2 and a standard deviation of 0.15. The distribution is truncated, with a lower bound of 0.7 and an upper bound of 2.75. Figure 5 shows the run thresholds as functions of bank heterogeneity Δ. As in the main text, heterogeneity is beneficial for stability of both weak and strong banks as long as it is not too large or small.

Figure 6 compares the cases in which the prior is informative and uninformative. In the former case, the parametrization is the same as above. In the latter case, the only difference is the prior distribution. In these two cases bank heterogeneities are not directly comparable, in the sense that the same values of the model primitives correspond to different Δ's. Because of that, we analyze how the run thresholds depend on the ex-post difference between asset performances of weak and strong banks η. In both models, an increase in η corresponds to an increase in bank heterogeneity Δ (panel (a) of Figure

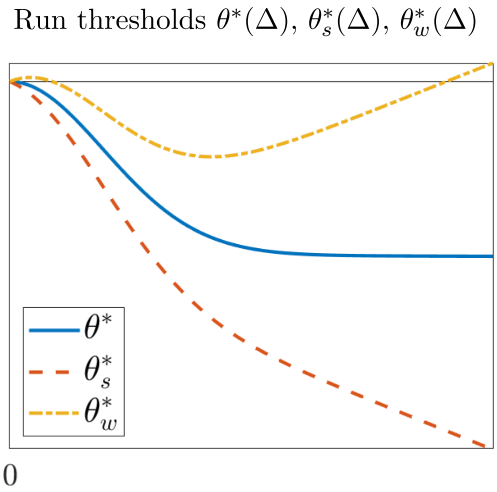

Figure 5: Run thresholds of strong (θ_s^*) and weak (θ_w^*) banks, and their average ($\theta^* = \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$) as functions of bank heterogeneity Δ. The prior is informative. See text for the numerical values used to get these graphs.

6). As Δ increases, the average fragility declines. Moreover, as long as Δ is not too large and there exist meaningful cross-bank strategic interactions, an increase in Δ tends to benefit both weak and strong banks. These results hold in both versions of the model (panels (b) and (c) of the same figure). One difference between the two models is that the run thresholds are overall lower when the prior is informative. This is because, for our choice of parameters, investors' prior information about the aggregate fundamental is plausible, and worse signals are required to trigger runs.

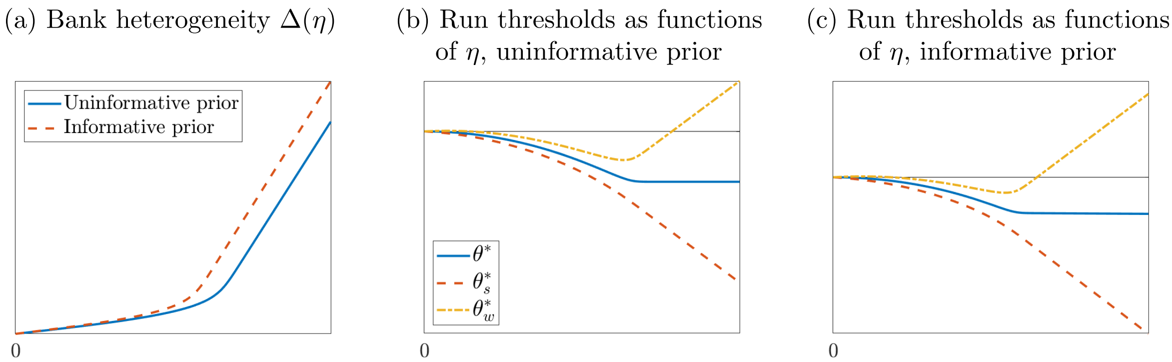

Figure 6: Panel (a): Bank heterogeneity Δ as a function of the size of bank-specific shock η for the cases in which prior about the aggregate fundamental is uninformative and informative. Panels (b) and (c): Run thresholds of strong (θ_s^*) and weak (θ_w^*) banks, and their average ($\theta^* = \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$) as functions of asset dispersion η for the cases of uninformative and informative priors, respectively. In panels (b) and (c), the Y-axes are aligned. See text for the numerical values used to get these graphs.

C.4 Bank failures

In the main model, we assume that bank i does not fail even if all "non-sleepy" investors withdraw their funds early, $m_i = \bar{m}$. In this appendix, we consider the model in which $\bar{m}$ is sufficiently high such that bank failures are possible. If bank i fails, it liquidates all assets and gets $\frac{z_i}{\lambda(m)}$ from outside investors. This amount is then split between all runners equally. The incremental payoff from staying (the analogue of Equation (2) in the main model) is

$$\pi = \begin{cases} \dfrac{z_i - m_i\lambda\left(m\right)}{1 - m_i} - 1 & \text{if } m_i\lambda\left(m\right) \leq z_i, \\ -\dfrac{z_i}{\lambda\left(m\right)m_i} & \text{if } m_i\lambda\left(m\right) > z_i. \end{cases}$$

As in Goldstein and Pauzner (2005), the possibility of failures creates a region of strategic substitution if $m_i\lambda\left(m\right) > z_i$. Furthermore, the possibility of failures implies that the within-bank and cross-bank strategic complementarities are no longer mutually reinforcing for all possible m_i and m. Specifically, we have $\frac{\partial \pi}{\partial m}\Big|_{m_i \to \left(\frac{z_i}{\lambda(m)}\right)^-} < 0$ and $\frac{\partial \pi}{\partial m}\Big|_{m_i \to \left(\frac{z_i}{\lambda(m)}\right)^+} > 0$. The proofs of Propositions 1 and 2 show that the complementarity reinforcement for all possible m_i and m is a sufficient condition under which bank heterogeneity is stabilizing, that is, $\frac{\partial \theta^*}{\partial \Delta} < 0$. Without this property, we no longer can prove this result analytically. However, as we show below numerically, even if the two complementarities are not mutually reinforcing everywhere, bank heterogeneity still tends to be associated with higher stability. Therefore, the complementarity reinforcement for all possible m_i and m is not a necessary condition for our main result.

Specifically, we focus on threshold equilibria and solve the following equations for θ^*

and Δ numerically.

$$\int_0^{x_s(\theta^*,\Delta)} \left(\frac{\theta^* + \eta - \lambda \left(\frac{\bar{m}}{2} x + \frac{\bar{m}}{2} F_\epsilon \left(\Delta + F_\epsilon^{-1}(x) \right) \right) \bar{m} x}{1 - \bar{m} x} - 1 \right) dx - \int_{x_s(\theta^*,\Delta)}^1 \frac{\theta^* + \eta}{\lambda \left(\frac{\bar{m}}{2} x + \frac{\bar{m}}{2} F_\epsilon \left(\Delta + F_\epsilon^{-1}(x) \right) \right) \bar{m} x} dx = 0, \tag{57}$$

$$\int_0^{x_w(\theta^*,\Delta)} \left(\frac{\theta^* - \eta - \lambda \left(\frac{\bar{m}}{2} x + \frac{\bar{m}}{2} F_\epsilon \left(-\Delta + F_\epsilon^{-1}(x) \right) \right) \bar{m} x}{1 - \bar{m} x} - 1 \right) dx - \int_{x_w(\theta^*,\Delta)}^1 \frac{\theta^* - \eta}{\lambda \left(\frac{\bar{m}}{2} x + \frac{\bar{m}}{2} F_\epsilon \left(-\Delta + F_\epsilon^{-1}(x) \right) \right) \bar{m} x} dx = 0, \tag{58}$$

where

$$x_s(\theta^*,\Delta) = \begin{cases} 1 & \text{if } \theta^* > \bar{m}\lambda \left(\frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon(\Delta + \bar{\epsilon}) \right) - \eta, \\ \hat{x}_s : \ \bar{m}\hat{x}_s \lambda \left(\frac{\bar{m}}{2}\hat{x}_s + \frac{\bar{m}}{2} F_\epsilon \left(\Delta + F_\epsilon^{-1}(\hat{x}_s) \right) \right) = \theta^* + \eta & \text{otherwise,} \end{cases}$$

and

$$x_w(\theta^*,\Delta) = \begin{cases} 1 & \text{if } \theta^* > \bar{m}\lambda \left(\frac{\bar{m}}{2} + \frac{\bar{m}}{2} F_\epsilon(-\Delta + \bar{\epsilon}) \right) + \eta, \\ \hat{x}_w : \ \bar{m}\hat{x}_w \lambda \left(\frac{\bar{m}}{2}\hat{x}_w + \frac{\bar{m}}{2} F_\epsilon \left(-\Delta + F_\epsilon^{-1}(\hat{x}_w) \right) \right) = \theta^* - \eta & \text{otherwise.} \end{cases}$$

Here $x_s(\theta^*,\Delta)$ and $x_w(\theta^*,\Delta)$ are the critical fractions of "non-sleepy" strong- and weak-bank runners below which strong and weak banks do not fail. For example, if θ^* and Δ are such that $x_s(\theta^*,\Delta) < 1$, it means that strong banks fail if a sufficiently large proportion of "non-sleepy" investors withdraw their funds early. If there is no solution for (57)-(58) in which $\Delta < \bar{\Delta}$, we solve (57) for θ_s^* and (58) for θ_w^* assuming $\Delta = \bar{\Delta}$. In that case, we define $\theta^* = \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$.

We consider the following parametrization: $\bar{m} \in \{0.55, 0.75, 0.90\}$; $\lambda(m) = 1 + m^2$;

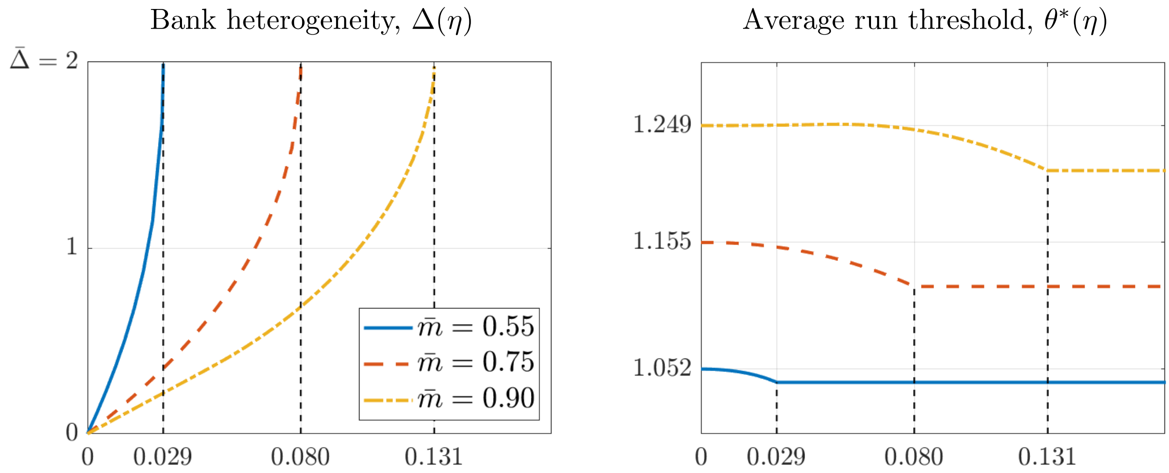

Figure 7: Bank heterogeneity $\Delta(\eta)$ (left panel) and average run threshold $\theta^*(\eta)$ (right panel) as functions of asset dispersion η. Parametrization: $\bar{m} \in \{0.55, 0.75, 0.90\}$; $\lambda(m) = 1 + m^2$; noise distribution is standard normal over $[-1, 1]$; η varies from 0 to 0.175.

noise distribution is standard normal over $[-1, 1]$; η varies from 0 to 0.175. Under our parametrization, there are no bank failures if $\bar{m} = 0.55$; however, bank failures are possible if $\bar{m} = 0.75$ and $\bar{m} = 0.90$. We plot $\Delta(\eta)$ and $\theta^*(\eta)$ in Figure 7. We find that an increase in η is associated with a higher bank heterogeneity Δ and lower average run threshold θ^* in all three cases (when there are nontrivial cross-bank strategic interactions, $\Delta < \bar{\Delta}$). Not surprisingly, we also find that the run problem is exacerbated if a larger fraction of investors are able to withdraw their funds early, that is, $\theta^*(\bar{m})$ is an increasing function.

C.5 Many bank types

Our baseline model assumes that bank-specific shocks take values η or $-\eta$ with equal probabilities. In this section, we show how our analyses can be extended to the case in which bank-specific shocks take $N \geq 2$ values.

The structure of the economy stays the same as in Section 2. The only difference is that bank-specific shock ζ_i can take $N \geq 2$ values, $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_N$, $\boldsymbol{\eta} = \{\eta_n\}_{n=1 \ldots N}$, with probabilities $\omega_1, \omega_2, \ldots, \omega_N$, respectively, where $\omega_n \in (0, 1)$ $\forall n \in \{1, \ldots, N\}$ and $\sum_{n=1}^{N} \omega_n = 1$. Without loss of generality, we assume that bank-specific shock is zero on average, $\sum_{n=1}^{N} \omega_n \eta_n = 0$.

Bank investors follow threshold strategies, that is, investors of a bank receiving a shock η_n withdraw early if their signals are below θ_n^* and do not do so otherwise. An indifference condition for an investor receiving a threshold signal θ_n^* is

$$\theta_n^* + \eta_n = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x} dx} \left(1 + \int_0^1 \lambda \left(\bar{m} \sum_{\tau=1}^{N} \omega_\tau F_\epsilon \left(\Delta_{n\tau} + F_\epsilon^{-1}(x) \right) \right) \frac{\bar{m}x}{1-\bar{m}x} dx \right), \quad (59)$$

where, as in the baseline model, signal noise is negligible, $\sigma \to 0$. $\Delta_{n\tau} = \lim_{\sigma \to 0} \frac{\theta_\tau^* - \theta_n^*}{\sigma}$ is the bank heterogeneity between banks receiving shocks η_n and η_τ, and $\Delta_{n\tau} \geq 0 \ \forall n < \tau$. A system of N equations (59) is a generalized version of Equations (7)-(8). In what follows, we focus on an interesting case of nontrivial cross-bank strategic interactions. That is, $\Delta_{n\tau} < \bar{\Delta} \ \forall n < \tau$. In this case, run thresholds are infinitely close to each other with the average run threshold given by

$$\theta^* \equiv \sum_{n=1}^{N} \omega_n \theta_n^* = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x} dx} \left(1 + \int_0^1 \sum_{n=1}^{N} \omega_n \lambda \left(\bar{m} \sum_{\tau=1}^{N} \omega_\tau F_\epsilon \left(\Delta_{n\tau} + F_\epsilon^{-1}(x) \right) \right) \frac{\bar{m}x}{1-\bar{m}x} dx \right). \quad (60)$$

The following proposition generalizes Proposition 1.

Proposition 8. *Suppose that $\Delta_{n\tau} < \bar{\Delta} \ \forall n < \tau$. In the limit of negligible information friction ($\sigma \to 0$), run thresholds of all banks are infinitely close to θ^*. Moreover, any change in bank heterogeneity that i) weakly increases $\Delta_{n\tau} \ \forall n < \tau$ and ii) strictly increases $\Delta_{n'\tau'}$ for some $n' < \tau'$ leads to a decline in θ^*.*

Proof. See Appendix C.5.1. □

Bank heterogeneity, described by pairwise distances between run thresholds $\Delta_{n\tau}$, is endogenous. As discussed in the previous sections, it depends on various primitives of the model, such as asset dispersion, information structure, and liquidity conditions. Given that this section generalizes the structure of bank-specific shocks, we focus on characterizing how differences in asset dispersion affect bank heterogeneity and hence the overall financial stability. The following proposition is a version of Proposition 3 in this more general model.

Proposition 9. *There exists an $\underline{\eta} > 0$ such that if $|\eta_n| < \underline{\eta} \ \forall n$, then in the limit of negligible information friction ($\sigma \to 0$), run thresholds of all banks are infinitely close to*

$\theta^*(\boldsymbol{\eta})$, which reaches a (local) maximum at $\boldsymbol{\eta} = \mathbf{0}$.

Proof. See Appendix C.5.2. □

If asset performances of all banks are identical, i.e. $\boldsymbol{\eta} = \mathbf{0}$, then all run thresholds are exactly the same and the system is homogeneous. By continuity, all run thresholds stay infinitely close to each other for any small change in $\boldsymbol{\eta}$. Pairwise heterogeneity terms $\Delta_{n\tau}$, however, adjust. In particular, if $\eta_n > \eta_\tau$, then $\Delta_{n\tau} > 0$. According to Proposition 9, such diversity in bank asset performances enhances stability of all banks.

As in the main model with two types of banks, there is a limit to which increasing diversity in bank-specific productivities is unequivocally stabilizing. In particular, if bank asset performance becomes sufficiently divergent—so that banks' behaviors in the asset market are fully decoupled and there are no strategic interactions across investors of different banks—further divergence hurts relatively weaker banks.

C.5.1 Proof of Proposition 8

Proof. From (59), the run threshold for a bank receiving a shock η_n is

$$\theta^* = -\eta_n + \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + \int_0^1 \lambda\left(\bar{m}\sum_{\tau=1}^{N}\omega_\tau F_\epsilon\left(\Delta_{1\tau} - \Delta_{1n} + F_\epsilon^{-1}(x)\right)\right)\frac{\bar{m}x}{1-\bar{m}x}dx\right),$$

where we impose that the run threshold θ_n^* is infinitely close to the average threshold θ^*. By definition, $\Delta_{n\tau} = \Delta_{1\tau} - \Delta_{1n}$. Notice that

$$\eta_n > \eta_\tau \Leftrightarrow \Delta_{1n} < \Delta_{1\tau} \text{ and } \eta_n = \eta_\tau \Leftrightarrow \Delta_{n\tau} = 0.$$

Recall that $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_N$. Heterogeneity in the financial system is fully described by $N-1$ variables $0 \leq \Delta_{12} \leq \Delta_{13} \leq \cdots \leq \Delta_{1N}$. A marginal increase in heterogeneity that (weakly) increases $\Delta_{n\tau}$ $\forall n < \tau$ corresponds to a change $0 \leq d\Delta_{12} \leq d\Delta_{13} \leq \cdots \leq d\Delta_{1N}$. Recall that the average fragility θ^* is given by Equation (60), which can be rewritten as

$$\theta^* = \frac{1}{\int_0^1 \frac{1}{1-\bar{m}x}dx}\left(1 + \int_0^1 \sum_n \omega_n \lambda\left(\bar{m}\sum_\tau \omega_\tau F_\epsilon\left(\Delta_{1\tau} - \Delta_{1n} + F_\epsilon^{-1}(x)\right)\right)\frac{\bar{m}x}{1-\bar{m}x}dx\right).$$

Denote

$$I\left(\Delta_{12},\ldots,\Delta_{1N}\right)=\int_0^1 f\left(x\right)\sum_n \omega_n\lambda\left(\bar{m}\sum_\tau \omega_\tau F_\epsilon\left(\Delta_{1\tau}-\Delta_{1n}+F_\epsilon^{-1}\left(x\right)\right)\right)dx,$$

where for brevity we denote $f(x)=\frac{\bar{m}x}{1-\bar{m}x}$. Differentiating $I\left(\Delta_{12},\ldots,\Delta_{1N}\right)$ with respect to Δ_{1k}, we get

$$\frac{\partial I}{\partial \Delta_{1k}}=\bar{m}\omega_k\sum_{n=1}^N \omega_n I_{nk},$$

where

$$I_{nk}=\int_0^1\left[f\left(x\right)-f\left(F_\epsilon\left(\Delta_{1k}-\Delta_{1n}+F_\epsilon^{-1}\left(x\right)\right)\right)\right]f_\epsilon\left(\Delta_{1k}-\Delta_{1n}+F_\epsilon^{-1}\left(x\right)\right)\times$$
$$\lambda'\left(\bar{m}\sum_\tau \omega_\tau F_\epsilon\left(\Delta_{1\tau}-\Delta_{1n}+F_\epsilon^{-1}\left(x\right)\right)\right)dx.$$

Clearly, $I_{kk}=0$ and $I_{1k}\leq 0\ \forall k\in\{2,\ldots,N\}$. Furthermore, if $k>n$, $\Delta_{1k}\geq\Delta_{1n}$ and hence $I_{nk}\leq 0$. By changing the variable of integration $x\to F_\epsilon\left(\Delta_{1k}-\Delta_{1n}+F_\epsilon^{-1}\left(x\right)\right)$, it is straightforward to derive that $I_{nk}=-I_{kn}$.

The impact of change in heterogeneity $0\leq d\Delta_{12}\leq\cdots\leq d\Delta_{1N}$ on $I\left(\Delta_{12},\ldots,\Delta_{1N}\right)$ is

$$dI=\sum_{k=2}^N\frac{\partial I}{\partial\Delta_{1k}}d\Delta_{1k}=\bar{m}\sum_{k=2}^N\sum_{n=1}^N\omega_k\omega_n I_{nk}d\Delta_{1k}=$$
$$\bar{m}\sum_{k=2}^N\omega_k\omega_1 I_{1k}d\Delta_{1k}+\bar{m}\sum_{k=2}^N\sum_{n=2}^N\omega_k\omega_n I_{nk}d\Delta_{1k}\leq 0.$$

The inequality holds because $I_{1k}\leq 0\ \forall k\in\{1,\ldots,N\}$ and $I_{nk}d\Delta_{1k}+I_{kn}d\Delta_{1n}=I_{nk}\left(d\Delta_{1k}-d\Delta_{1n}\right)\leq 0$. The latter is true because if $k>n$, $I_{nk}\leq 0$ and $d\Delta_{1k}\geq d\Delta_{1n}$, and if $k<n$, $I_{nk}\geq 0$ and $d\Delta_{1k}\leq d\Delta_{1n}$.

Therefore, a marginal change in heterogeneity that does not reduce pairwise heterogeneities $\Delta_{n\tau}\ \forall n<\tau$ and increases $\Delta_{n'\tau'}$ for some $n'<\tau'$ leads to a lower $I\left(\Delta_{12},\ldots,\Delta_{1N}\right)$ and hence a lower θ^*. □

C.5.2 Proof of Proposition 9

Proof. If all bank-specific productivities are identical, i.e. $\boldsymbol{\eta}=\mathbf{0}$, the banking system is homogeneous. Therefore, $\Delta_{n\tau}=0\ \forall n,\tau$, and all investors run if their signals are below

$\theta^*(\mathbf{0})$. By continuity of $\lambda(\cdot)$ and $F_\epsilon(\cdot)$, there must exist an $\underline{\eta} > 0$ such that if $|\eta_n| < \underline{\eta}\ \forall n$, then investors of all banks share the same run threshold $\theta^*(\boldsymbol{\eta})$.

Consider an $\boldsymbol{\eta} \neq \mathbf{0}$ with $|\eta_n| < \underline{\eta}\ \forall n$. For such $\boldsymbol{\eta}$, without loss of generality we can write $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_N$, with at least one inequality being strict. Therefore, $0 \leq \Delta_{12} \leq \Delta_{13} \leq \cdots \leq \Delta_{1N}$, with at least one inequality being strict. By Proposition 8, $\theta^*(\boldsymbol{\eta}) < \theta^*(\mathbf{0})$. □

C.6 Cross-bank deposit flows

In our model, bank investors who withdraw their funds early do not redeposit them to other banks. As such, our paper directly applies to situations in which capital outflows form the entire banking system. In practice, some of weak-bank depositors may run to strong banks, as was the case, for example, during the 2023 banking crisis. Strong banks then can use new deposits to acquire assets liquidated by weak banks. While an in-depth modeling of such cross-bank runs is outside the scope of this paper, a simple way to accommodate such a possibility in our framework is as follows.

The model extends the one described in Section 2. The timeline of the extended model is in Figure 8. Specifically, suppose that at $t = 1$ the total mass of runs is m, and suppose that a mass $n \leq m$ of runners redeposit their money at a deposit rate of one right upon withdrawal. That is, they expect to receive a payoff of one at $t = 2$ for each unit of investment. They are willing to accept such a deposit rate at $t = 1$ because they do not discount the future.

To repay runners, banks sell their long-term assets to outside investors. However, unlike our main model, banks can use redeposited money to immediately acquire some of the liquidated assets from outsiders. We assume that outside investors do not incur any losses if they have to hold assets for only a short period of time. However, if they have to raise L to acquire long-term assets and hold them to maturity, they incur a cost $g(L) \geq L$.

We assume that banks are efficient users of liquidated assets. Then bank i acquiring one unit of bank j's long-term assets expects to get a gross return z_j. The price bank i is willing to pay for each unit of bank j's assets is z_j. Note that in this simple model, it is not important which banks receive deposits from runners at $t = 1$—the key is that they stay with banks who are efficient users of other banks' assets.

In this setting, if the mass of runs in the economy at $t = 1$ is m, the total amount of liquidity outside investors need to raise to acquire held-to-maturity long-term assets is $[m - n]^+ = \max\{m - n, 0\}$. The incremental payoff from staying for bank investors is therefore

$$\pi(z_i, m_i, m) = \frac{z_i - m_i \lambda([m - n]^+)}{1 - m_i} - 1.$$

Notice that an increase in n affects run incentives in the same way as liquidity injections considered in Section 4.3. In particular, larger the mass of runners who redeposit their funds to other banks, less severe the fire-sale problem is, and more stable the financial system is (by Proposition 5, $\frac{d\theta^*}{dn} < 0$).

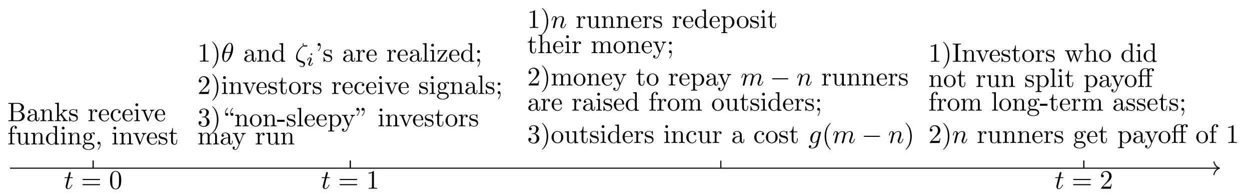

Figure 8: Timeline

Of course, this simple model has several important limitations. For example, we do not model how investors who withdraw their funds at $t = 1$ decide whether to redeposit their money. In practice, their decisions are likely driven by their beliefs about redepositing behavior of other investors and information about economic fundamentals they receive after their initial withdrawal decisions. Furthermore, we assume that if bank i has to liquidate its assets but sells them to other banks, it does not incur any fire-sale losses. In practice, such a bank needs to sell quickly which likely weakens its bargaining position. In that case, the price bank i will receive is likely below z_i even if the acquiring bank is an efficient user of bank i's assets. We believe that it would be interesting to incorporate these realistic features in a more detailed model to study cross-bank deposit flows in depth.

C.7 Endogenous liquidity buffers

In this appendix, we consider a model in which banks pick their liquidity buffers endogenously. As in Section 4.4, we assume that the fire-sale discount function is weakly convex, that is, $\lambda''(\cdot) \geq 0$. To keep the analysis concise, we assume that the support of the noise is unbounded such that $\underline{\epsilon} = -\infty$ and $\bar{\epsilon} = \infty$.

Bank i chooses the liquidity buffer size l_i before observing its bank-specific shock and before any signals about the aggregate fundamentals are realized. The bank's objective is to maximize its total expected payout to investors (see Equation (65) below).

We focus on symmetric threshold equilibria so that bank i expects all other banks to pick the liquidity buffer size l (recall that liquidity buffer sizes are chosen before any shocks are realized, and all banks are identical at that point). After bank-specific shocks are realized and bank investors receive signals about the aggregate fundamental, they may decide to withdraw their funds early. If a fraction x of "non-sleepy" investors run on bank i, then the total mass of runs in the economy is $m^i_{tot}(x) = \frac{\bar{m}}{2} F_\epsilon\left(\frac{\theta^*_s - \theta^*_i}{\sigma} + F^{-1}_\epsilon(x)\right) + \frac{\bar{m}}{2} F_\epsilon\left(\frac{\theta^*_w - \theta^*_i}{\sigma} + F^{-1}_\epsilon(x)\right)$, where θ^*_s and θ^*_w are the run thresholds for banks that choose l as their liquidity buffer size and that have bank-specific shocks of η and $-\eta$, respectively, and θ^*_i is the run threshold for bank i. Note that, because banks are atomistic, the total mass of runs $m^i_{tot}(x)$ does not depend on the mass of runs on bank i.

The incremental payoff from staying for investors of bank i is

$$\begin{aligned}\pi_i(x) = & \frac{\left(\theta^*_i + \zeta_i - \sigma F^{-1}_\epsilon(x)\right)(1 - l_i)}{1 - \bar{m}x} + \frac{l_i - \bar{m}x}{1 - \bar{m}x} \mathbb{1}\{l_i \geq \bar{m}x\} - \\ & \frac{\lambda\left([m^i_{tot}(x) - l]^+\right)(\bar{m}x - l_i)}{1 - \bar{m}x} \mathbb{1}\{l_i < \bar{m}x\} - 1.\end{aligned} \tag{61}$$

If the liquidity buffer size of bank i exceeds the mass of runs on this bank, $l_i \geq \bar{m}x$, then it does not need to liquidate any assets. Furthermore, if the total mass of runs in the economy does not exceed the aggregate liquidity buffer, $m^i_{tot}(x) \leq l$, then bank i does not face any discount even if it is forced to liquidate its assets. In that case, other banks have enough liquidity to acquire assets liquidated by bank i.

The following proposition states the main result of this appendix. In particular, it compares marginal benefits of increasing the size of a liquidity buffer from the perspective of an individual bank and from the perspective of the planner. The planner is assumed to maximize aggregate output by choosing l that is common to all banks (see Equation (67) below).

Proposition 10. *Suppose that the long-run technology is more productive than liquid assets, $\int \theta dF_\theta(\theta) > 1$. In a symmetric equilibrium, $l_i = l$, $\theta_i^* = \theta_s^*$ if $\zeta_i = \eta$, and $\theta_i^* = \theta_w^*$ if $\zeta_i = -\eta$. There exists $\bar{\sigma}$ such that if $\sigma < \bar{\sigma}$ then the liquidity buffer chosen by banks l_{eq}^* does not exceed the liquidity buffer chosen by the planner l_{pl}^*, $0 \leq l_{eq}^* \leq l_{pl}^* < \bar{m}$. Furthermore, $l_{eq}^* = l_{pl}^*$ only if $l_{eq}^* = l_{pl}^* = 0$.*

Proposition 10 implies that individual banks choose a lower liquidity buffer relative to the planner (unless the planner chooses $l = 0$, in which case individual banks also choose $l = 0$). An individual bank internalizes that a higher l means that a smaller fraction of capital is invested in long-term assets. It also internalizes that a higher l means that its investors are less likely to run and that it will have to sell fewer assets at a fire-sale discount if the run does happen. However, it does not internalize how its actions affect stability of other banks. Specifically, if an individual bank i increases its liquidity buffer, it marginally reduces fire-sale discounts for all other banks. Their investors are then less likely to run, which is beneficial for bank i as her investors then expect a smaller fire-sale discount. As a result, individual banks underinvest in liquid assets.[28]

Proof. **Part I: Individual bank**

The run threshold for bank i is implicitly determined by $\int_0^1 \pi_i(x)dx = 0$, where $\pi_i(x)$ is given by (61). Specifically,[29]

$$\theta_i^* + \zeta_i = \frac{\int_{\frac{l_i}{\bar{m}}}^1 \frac{\lambda\left(\left[m_{tot}^i(x) - l\right]^+\right)(\bar{m}x - l_i)}{1 - \bar{m}x} dx + 1 - \int_0^{\frac{l_i}{\bar{m}}} \frac{l_i - \bar{m}x}{1 - \bar{m}x} dx}{\int_0^1 \frac{1 - l_i}{1 - \bar{m}x} dx} + \frac{\sigma \int_0^1 \frac{F_\epsilon^{-1}(x)}{1 - \bar{m}x} dx}{\int_0^1 \frac{dx}{1 - \bar{m}x}}.$$

[28] In our model, markets are incomplete, and so the pecuniary externality can have real consequences (Greenwald and Stiglitz, 1986).

[29] All the derivations in this proof are made under the assumption that signals are much more precise than any prior information that investors have about the aggregate fundamental θ. As we are going to focus on the case of small σ, this is without loss of generality.

Denote the right-hand side of the above expression by T. Then

$$\frac{d\theta_i^*}{dl_i} = \frac{\frac{\partial T}{\partial l_i}}{1 - \frac{\partial T}{\partial \theta_i^*}}, \tag{62}$$

where

$$\frac{\partial T}{\partial l_i} = \frac{-\int_{\frac{l_i}{\bar{m}}}^1 \left(\lambda\left([m_{tot}^i(x) - l]^+\right) - 1\right) dx}{(1-l_i)^2 \left(\int_0^1 \frac{1}{1-\bar{m}x} dx\right)} \leq 0,$$

$$\frac{\partial T}{\partial \theta_i^*} = -\frac{\frac{\bar{m}}{2}\frac{1}{\sigma}}{\int_0^1 \frac{1-l_i}{1-\bar{m}x} dx} \int_{\frac{l_i}{\bar{m}}}^1 \frac{\lambda'\left(m_{tot}^i(x) - l\right)\left(\bar{m}x - l_i\right)}{1-\bar{m}x} \times$$
$$\left(f_\epsilon\left(\frac{\theta_s^* - \theta_i^*}{\sigma} + F_\epsilon^{-1}(x)\right) + f_\epsilon\left(\frac{\theta_w^* - \theta_i^*}{\sigma} + F_\epsilon^{-1}(x)\right)\right) \mathbb{1}\left\{m_{tot}^i(x) - l \geq 0\right\} dx \leq 0.$$

Clearly, $\frac{d\theta_i^*}{dl_i} \leq 0$.

Given the run threshold θ_i^* and conditional on the aggregate state θ, the mass of runs on bank i is $\bar{m}F_\epsilon\left(\frac{\theta_i^*-\theta}{\sigma}\right)$. The total mass of runs in the economy is $m_{tot}(\theta) = \frac{\bar{m}}{2}F_\epsilon\left(\frac{\theta_s^*-\theta}{\sigma}\right) + \frac{\bar{m}}{2}F_\epsilon\left(\frac{\theta_w^*-\theta}{\sigma}\right)$. Conditional on the bank-specific shock ζ_i, the expected payout of bank i to its investors is

$$y_i(\zeta_i) =$$
$$\int \bar{m}F_\epsilon\left(\frac{\theta_i^* - \theta}{\sigma}\right) dF_\theta(\theta) + \int_{\theta_i^* - \sigma F_\epsilon^{-1}\left(\frac{l_i}{\bar{m}}\right)}^\infty \left[(\theta + \zeta_i)(1 - l_i) + l_i - \bar{m}F_\epsilon\left(\frac{\theta_i^* - \theta}{\sigma}\right)\right] dF_\theta(\theta) +$$
$$\int_{-\infty}^{\theta_i^* - \sigma F_\epsilon^{-1}\left(\frac{l_i}{\bar{m}}\right)} \left[(\theta + \zeta_i)(1 - l_i) - \lambda\left([m_{tot}(\theta) - l]^+\right)\left(\bar{m}F_\epsilon\left(\frac{\theta_i^* - \theta}{\sigma}\right) - l_i\right)\right] dF_\theta(\theta).$$

Taking derivatives of y_i with respect to l_i and θ_i^*, we get

$$\frac{\partial y_i}{\partial l_i} = -\int \theta dF_\theta(\theta) - \zeta_i + \int_{\theta_i^* - \sigma F_\epsilon^{-1}\left(\frac{l_i}{\bar{m}}\right)}^\infty dF_\theta(\theta) + \int_{-\infty}^{\theta_i^* - \sigma F_\epsilon^{-1}\left(\frac{l_i}{\bar{m}}\right)} \lambda\left([m_{tot}(\theta) - l]^+\right) dF_\theta(\theta), \tag{63}$$

$$\frac{\partial y_i}{\partial \theta_i^*} = -\int_{-\infty}^{\theta_i^* - \sigma F_\epsilon^{-1}\left(\frac{l_i}{\bar{m}}\right)} \left(\lambda\left([m_{tot}(\theta) - l]^+\right) - 1\right) \bar{m}\frac{1}{\sigma} f_\epsilon\left(\frac{\theta_i^* - \theta}{\sigma}\right) dF_\theta(\theta). \tag{64}$$

Bank i chooses $l_i \geq 0$ to maximize

$$y_i^{tot}(l_i) = \frac{1}{2}y_i(\eta) + \frac{1}{2}y_i(-\eta). \tag{65}$$

Using (62), (63), and (64), it is straightforward to derive the full derivative of (65) with respect to l_i. The first-order conditions are

$$0 = \gamma_i + \frac{1}{2} \sum_{k \in \{s,w\}} \left(\frac{\partial y_i}{\partial l_i} + \frac{\partial y_i}{\partial \theta_i^*} \frac{d\theta_i^*}{dl_i} \right) \Bigg|_{l_i = l, \theta_i^* = \theta_k^*, \zeta_i = \zeta_k}, \tag{66}$$

$$0 = \gamma_i l_i,$$

where $\zeta_k = \eta$ if $k = s$ and $\zeta_k = -\eta$ if $k = w$. In the expressions above, $\gamma_i \geq 0$ is the Lagrange multiplier associated with the $l_i \geq 0$ constraint.

Finally, from (63) it is clear that $\frac{\partial y_i}{\partial l_i} \overset{l_i \to \bar{m}}{\to} 1 - \int \theta dF_\theta(\theta) - \zeta_i$, and from (64) $\frac{\partial y_i}{\partial \theta_i^*} \overset{l_i \to \bar{m}}{\to} 0$. Therefore, $\frac{dy_i^{tot}}{dl_i} \overset{l_i \to \bar{m}}{\to} 1 - \int \theta dF_\theta(\theta)$. If the long-run technology is more productive than liquid assets, i.e. $\int \theta dF_\theta(\theta) > 1$, then $l_i^* = l_{eq}^* < \bar{m}$.

Part II: Planner

The run thresholds for strong and weak banks are, respectively, θ_s^* and θ_w^*. Given these run thresholds and conditional on the aggregate state θ, the masses of runs on each strong and weak bank are, respectively, $m_s(\theta) = \bar{m} F_\epsilon \left(\frac{\theta_s^* - \theta}{\sigma} \right)$ and $m_w(\theta) = \bar{m} F_\epsilon \left(\frac{\theta_w^* - \theta}{\sigma} \right)$, with $m_s(\theta) < m_w(\theta)$. The total mass of runs in the economy is then $m_{tot}(\theta) = \frac{1}{2} m_s(\theta) + \frac{1}{2} m_w(\theta)$. The planner chooses l to maximize expected output Y given below.

$$Y = Y_1 + Y_2 + Y_3 + Y_4, \tag{67}$$

where Y_1 is the total payoff to runners,

$$Y_1 = \frac{1}{2} \int m_s(\theta) \, dF_\theta(\theta) + \frac{1}{2} \int m_w(\theta) \, dF_\theta(\theta),$$

Y_2 is the total payoff to strong-bank investors who do not run,

$$Y_2 = \frac{1}{2} \int (1 - m_s(\theta)) \left[\frac{(\theta + \eta)(1 - l)}{1 - m_s(\theta)} + \frac{l - m_s(\theta)}{1 - m_s(\theta)} \mathbb{1}\{m_s(\theta) \leq l\} - \right.$$
$$\left. \frac{\lambda (m_{tot}(\theta) - l)(m_s(\theta) - l)}{1 - m_s(\theta)} \mathbb{1}\{m_s(\theta) > l\} \right] dF_\theta(\theta),$$

Y_3 is the total payoff to weak-bank investors who do not run,

$$Y_3 = \frac{1}{2}\int (1 - m_w(\theta)) \left[\frac{(\theta - \eta)(1 - l)}{1 - m_w(\theta)} + \frac{l - m_w(\theta)}{1 - m_w(\theta)} \mathbb{1}\{m_{tot}(\theta) \leq l\} - \right.$$
$$\left. \frac{\lambda (m_{tot}(\theta) - l)(m_w(\theta) - l)}{1 - m_w(\theta)} \mathbb{1}\{m_{tot}(\theta) > l\} \right] dF_\theta(\theta),$$

and Y_4 is the total payoff to outsiders,

$$Y_4 = \int [(m_{tot}(\theta) - l)\lambda(m_{tot}(\theta) - l) - g(m_{tot}(\theta) - l)] \mathbb{1}\{m_{tot}(\theta) > l\} dF_\theta(\theta).$$

The derivative of Y with respect to l is

$$\frac{dY}{dl} = \frac{\partial Y}{\partial l} + \frac{\partial Y}{\partial \theta_s^*}\frac{d\theta_s^*}{dl} + \frac{\partial Y}{\partial \theta_w^*}\frac{d\theta_w^*}{dl}. \tag{68}$$

Taking partial derivatives of Y with respect to l, θ_s^* and θ_w^*, we get

$$\begin{aligned} \frac{\partial Y}{\partial l} = & - \int \theta dF_\theta(\theta) + \frac{1}{2}\int [\mathbb{1}\{m_s(\theta) \leq l\} + \lambda(m_{tot}(\theta) - l)\mathbb{1}\{m_s(\theta) > l\}] dF_\theta(\theta) + \\ & \frac{1}{2}\int [\mathbb{1}\{m_{tot}(\theta) \leq l\} + \lambda(m_{tot}(\theta) - l)\mathbb{1}\{m_{tot}(\theta) > l\}] dF_\theta(\theta) + \\ & \frac{1}{2}\int \lambda'(m_{tot}(\theta) - l)(l - m_s(\theta))\mathbb{1}\{m_s(\theta) \leq l\}\mathbb{1}\{m_{tot}(\theta) > l\} dF_\theta(\theta), \end{aligned} \tag{69}$$

and

$$\begin{aligned} \frac{\partial Y}{\partial \theta_s^*} = & -\frac{1}{2}\int \bar{m}\frac{1}{\sigma}f_\epsilon\left(\frac{\theta_s^* - \theta}{\sigma}\right)[\lambda(m_{tot}(\theta) - l) - 1]\mathbb{1}\{m_s(\theta) > l\} dF_\theta(\theta) - \\ & \frac{1}{2}\int \frac{\bar{m}}{2}\frac{1}{\sigma}f_\epsilon\left(\frac{\theta_s^* - \theta}{\sigma}\right)\lambda'(m_{tot}(\theta) - l)(l - m_s(\theta))\mathbb{1}\{m_s(\theta) \leq l\}\mathbb{1}\{m_{tot}(\theta) > l\} dF_\theta(\theta), \end{aligned} \tag{70}$$

and

$$\begin{aligned} \frac{\partial Y}{\partial \theta_w^*} = & -\frac{1}{2}\int \bar{m}\frac{1}{\sigma}f_\epsilon\left(\frac{\theta_w^* - \theta}{\sigma}\right)[\lambda(m_{tot}(\theta) - l) - 1]\mathbb{1}\{m_{tot}(\theta) > l\} dF_\theta(\theta) - \\ & \frac{1}{2}\int \frac{\bar{m}}{2}\frac{1}{\sigma}f_\epsilon\left(\frac{\theta_w^* - \theta}{\sigma}\right)\lambda'(m_{tot}(\theta) - l)(l - m_s(\theta))\mathbb{1}\{m_s(\theta) \leq l\}\mathbb{1}\{m_{tot}(\theta) > l\} dF_\theta(\theta). \end{aligned} \tag{71}$$

Finally, we need to derive how the run thresholds θ_s^* and θ_w^* change with l. The run thresholds are given by

$$\theta_s^* + \eta = \frac{\int_{\frac{l}{\bar{m}}}^1 \frac{\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x}dx + 1 - \int_0^{\frac{l}{\bar{m}}} \frac{l-\bar{m}x}{1-\bar{m}x}dx}{\int_0^1 \frac{1-l}{1-\bar{m}x}dx} + \frac{\sigma\int_0^1 \frac{F_\epsilon^{-1}(x)}{1-\bar{m}x}dx}{\int_0^1 \frac{dx}{1-\bar{m}x}}, \tag{72}$$

$$\theta_w^* - \eta = \frac{\int_{x_w(\Delta,l)}^1 \frac{\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x}dx + 1 - \int_0^{x_w(\Delta,l)} \frac{l-\bar{m}x}{1-\bar{m}x}dx}{\int_0^1 \frac{1-l}{1-\bar{m}x}dx} + \frac{\sigma\int_0^1 \frac{F_\epsilon^{-1}(x)}{1-\bar{m}x}dx}{\int_0^1 \frac{dx}{1-\bar{m}x}}, \tag{73}$$

where $\Delta = \frac{\theta_w^* - \theta_s^*}{\sigma}$ and $x_w(\Delta, l)$ solves $\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) = l$. Note that there always exists a unique solution $x_w(\Delta, l) \in [0,1]$ if $l \in [0, \bar{m}]$.

We have

$$\theta_s^* = \theta^* - \frac{1}{2}\sigma\Delta,$$
$$\theta_w^* = \theta^* + \frac{1}{2}\sigma\Delta,$$

where $\theta^* = \frac{1}{2}\theta_s^* + \frac{1}{2}\theta_w^*$. Following the same steps as in the proof of Lemma 7, we can show that $\frac{\partial\theta^*}{\partial l} < 0$ and $\frac{\partial\theta^*}{\partial\Delta} < 0$.

From (72) and (73),

$$\sigma\Delta = 2\eta + \frac{1}{\int_0^1 \frac{1-l}{1-\bar{m}x}dx}\left(\int_{x_w(\Delta,l)}^1 \frac{\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x}dx - \right.$$
$$\left.\int_0^{x_w(\Delta,l)} \frac{l-\bar{m}x}{1-\bar{m}x}dx - \int_{\frac{l}{\bar{m}}}^1 \frac{\lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\Delta + F_\epsilon^{-1}(x)\right) - l\right)(\bar{m}x - l)}{1-\bar{m}x}dx + \int_0^{\frac{l}{\bar{m}}} \frac{l-\bar{m}x}{1-\bar{m}x}dx\right).$$

Denote the right-hand side of the expression above by B. Then

$$\frac{d\Delta}{dl} = -\frac{\frac{\partial B}{\partial l}}{\frac{\partial B}{\partial\Delta} - \sigma}.$$

Following the same steps as in the proof of Lemma 5, we can establish that $\frac{\partial B}{\partial l} < 0$ and $\frac{\partial B}{\partial\Delta} > 0$. Therefore, for a sufficiently small σ, $\frac{d\Delta}{dl} > 0$.

Therefore, we have

$$\frac{d\theta_s^*}{dl} = \frac{\partial\theta^*}{\partial l} + \frac{\partial\theta^*}{\partial\Delta}\frac{d\Delta}{dl} - \frac{1}{2}\sigma\frac{d\Delta}{dl},$$
$$\frac{d\theta_w^*}{dl} = \frac{\partial\theta^*}{\partial l} + \frac{\partial\theta^*}{\partial\Delta}\frac{d\Delta}{dl} + \frac{1}{2}\sigma\frac{d\Delta}{dl}.$$

For a sufficiently small σ, we therefore have $\frac{d\theta_s^*}{dl} < 0$ and $\frac{d\theta_w^*}{dl} < 0$.

Using the expressions above, we can write the first-order conditions are for the planner's problem of the l choice as

$$0 = \Gamma + \frac{\partial Y}{\partial l} + \frac{\partial Y}{\partial\theta_s^*}\frac{d\theta_s^*}{dl} + \frac{\partial Y}{\partial\theta_w^*}\frac{d\theta_w^*}{dl}, \tag{74}$$
$$0 = \Gamma l,$$

where $\Gamma \geq 0$ is the Lagrange multiplier associated with the $l \geq 0$ constraint.

From (68) it is clear that $\frac{\partial Y}{\partial l} \overset{l\to\bar{m}}{\to} 1 - \int \theta dF_\theta(\theta)$, from (70) we have $\frac{\partial Y}{\partial\theta_s^*} \overset{l\to\bar{m}}{\to} 0$, and from (71) we have $\frac{\partial Y}{\partial\theta_w^*} \overset{l\to\bar{m}}{\to} 0$. Therefore, $\frac{dY}{dl} \overset{l\to\bar{m}}{\to} 1 - \int \theta dF_\theta(\theta)$. If the long-run technology is more productive than liquid assets, i.e. $\int \theta dF_\theta(\theta) > 1$, then $l^* = l_{pl}^* < \bar{m}$.

Part III: Comparison

The final step is to show that either $l_{pl}^* > l_{eq}^*$ or $l_{pl}^* = l_{eq}^* = 0$. To do so, we are going to compare the right-hand sides of (66) and (74). Specifically, we are going to show that for a sufficiently small σ,

$$\frac{\partial Y}{\partial l} + \frac{\partial Y}{\partial\theta_s^*}\frac{d\theta_s^*}{dl} + \frac{\partial Y}{\partial\theta_w^*}\frac{d\theta_w^*}{dl} > \frac{1}{2}\sum_{k\in\{s,w\}}\left(\frac{\partial y_i}{\partial l_i} + \frac{\partial y_i}{\partial\theta_i^*}\frac{d\theta_i^*}{dl_i}\right)\Bigg|_{l_i=l,\theta_i^*=\theta_k^*,\zeta_i=\zeta_k}$$

for any given value of l.

Using (63) and (69), we find

$$\frac{\partial Y}{\partial l} - \frac{1}{2}\sum_{k\in\{s,w\}}\frac{\partial y_i}{\partial l_i}\Bigg|_{l_i=l,\theta_i^*=\theta_k^*,\zeta_i=\zeta_k} =$$
$$\frac{1}{2}\int \lambda'\left(m_{tot}\left(\theta\right)-l\right)\left(l-m_s\left(\theta\right)\right)\mathbb{1}\left\{m_s\left(\theta\right)\leq l\right\}\mathbb{1}\left\{m_{tot}\left(\theta\right)>l\right\}dF_\theta\left(\theta\right) > 0.$$

Similarly, using (63), (70), and (71), one can show that $\frac{\partial Y}{\partial\theta_s^*} < \frac{1}{2}\frac{\partial y_i}{\partial\theta_i^*}\Big|_{l_i=l,\theta_i^*=\theta_s^*,\zeta_i=\eta} < 0$ and $\frac{\partial Y}{\partial\theta_w^*} < \frac{1}{2}\frac{\partial y_i}{\partial\theta_i^*}\Big|_{l_i=l,\theta_i^*=\theta_w^*,\zeta_i=-\eta} < 0$. Finally, from (62), $\frac{d\theta_i^*}{dl_i} \to 0$ if $\sigma \to 0$ because

$\frac{\partial T}{\partial \theta_i^*} \to -\infty$, while from the planner's perspective, $\frac{d\theta_s^*}{dl}$ and $\frac{d\theta_w^*}{dl}$ are strictly negative even if σ is small.

All in all, the marginal benefit of increasing the size of the liquidity buffer is always higher from the planner's perspective than from individual bank's perspective. If the planner finds it optimal to set l to zero—for example, if the long-run technology is much more productive than liquid assets, $\int \theta dF_\theta(\theta) \gg 1$—then individual banks also do so. If the planner sets $l = l_{pl}^* > 0$, then $l_{eq}^* < l_{pl}^*$. $\square$

D Global games proofs

In this appendix, we prove that our baseline model features a unique threshold equilibrium. Our proof is relatively standard and based on Morris and Shin (2003).

For investors of bank i, the net benefit of not withdrawing funds early is

$$\pi\left(\theta+\zeta_i, m_i, m\right) = \frac{\theta+\zeta_i-\lambda(m)m_i}{1-m_i} - 1 = \frac{\theta+\zeta_i-1-\left(\lambda(m)-1\right)m_i}{1-m_i}. \tag{75}$$

Here m_i is the mass of runners on bank i, and $m = \int m_i di$ is the total mass of runners in the whole economy. Idiosyncratic productivity ζ_i takes values η with probability $\frac{1}{2}$ and $-\eta$ with probability $\frac{1}{2}$.

Investor l of bank i receives a signal about the aggregate fundamental θ,

$$s_{il} = \theta + \sigma\epsilon_{il}.$$

Noise terms ϵ_{il} are independent across investors and have identical cumulative distribution function $F_\epsilon(\cdot)$ that is increasing on the support $[\underline{\epsilon}, \bar{\epsilon}]$, $-\infty \le \underline{\epsilon} \le \bar{\epsilon} \le \infty$. As in the main text, $\mathbb{E}\epsilon_{il} = 0$ and $\mathbb{V}\epsilon_{il} = 1$. In this appendix, we assume that noise is small, that is, $\sigma \to 0$. Using continuity arguments as in Morris and Shin (2003), one can easily show that our results can be extended to the case in which signals are sufficiently, but not necessarily infinitely, more precise than the prior.

Notice that the payoff function (75) is not always decreasing in m_i, that is, it does not always imply strategic complementarities across investors of the same bank. However, as we show in what follows, it features a single-crossing property. As a result, under a set of standard assumptions outlined below, the game features a unique threshold equilibrium. However, as in Morris and Shin (2003), we cannot rule out existence of non-threshold equilibria.

Since we focus on threshold equilibria, we can rewrite the payoff function as

$$\Pi(z_i, m_i, \hat{\Delta}_i) \equiv \pi\left(z_i, m_i, \frac{m_i}{2} + \frac{\bar{m}}{2} F_\epsilon\left(\hat{\Delta}_i + F_\epsilon^{-1}\left(\frac{m_i}{\bar{m}}\right)\right)\right).$$

Here $\hat{\Delta}_i$ represents a signed distance between run thresholds normalized by σ. Specifically, for strong banks $\hat{\Delta}_s = \lim_{\sigma\to 0} \frac{\theta_w^* - \theta_s^*}{\sigma}$, and for weak banks $\hat{\Delta}_w = -\hat{\Delta}_s$. With this alternative payoff expression, we can define single-crossing property as follows.

Definition 3. *A payoff function $\pi(z_i, m_i, m)$ satisfies a single-crossing property if for any $z_i \in [\underline{\theta} - \eta, \bar{\theta} + \eta]$ and any $\hat{\Delta}_i \in [\underline{\epsilon} - \bar{\epsilon}, \bar{\epsilon} - \underline{\epsilon}]$, there exists at most one $m_i^* \in (0, \bar{m})$ such that $\Pi(z_i, m_i, \hat{\Delta}_i)$ switches sign, i.e., $\Pi \gtreqless 0$ if $m_i \lesseqgtr m_i^*$.*

It is straightforward to verify that the payoff function in our main model, given by (75), satisfies the single-crossing property.

As we show in the proof of Proposition 11 below, the unique threshold equilibrium result can be generalized to any payoff function that satisfies the single-crossing property. We make the following standard assumptions.

Assumption 2. *(Dominance regions) There exist θ^{LDR} and θ^{UDR} such that $\pi(\theta^{LDR} + \eta, 0, 0) < 0$ and $\pi(\theta^{UDR} - \eta, \bar{m}, \bar{m}) > 0$.*

Assumption 3. *(Monotone likelihood ratio property) Probability density function of noise $f_\epsilon(\cdot)$ is such that $\frac{f_\epsilon(\xi^H - \epsilon)}{f_\epsilon(\xi^L - \epsilon)}$ increases in ϵ for any $\xi^H > \xi^L$.*

Proposition 11. *Given any payoff function $\pi(z_i, m_i, m)$ that increases in z_i, decreases in m and satisfies the single-crossing property, there exists a unique threshold equilibrium in which investors of strong and weak banks withdraw early if their signals are below θ_s^* and θ_w^*, respectively, and do not withdraw early otherwise.*

Proof. Define $\pi^s(s, k_s, k_w)$ as the net benefit of not running on her bank for a strong-bank investor that observes signal s and believes that investors of strong and weak banks run if their signals are below k_s and k_w, respectively. Define $\pi^w(s, k_w, k_s)$ in an analogous way but for a weak-bank investor.

$$\pi^s(s, k_s, k_w) = \frac{\int_{s-\sigma\bar{\epsilon}}^{s-\sigma\underline{\epsilon}} \pi\left(\theta + \eta, \bar{m}F_\epsilon\left(\frac{k_s-\theta}{\sigma}\right), \frac{\bar{m}}{2}F_\epsilon\left(\frac{k_s-\theta}{\sigma}\right) + \frac{\bar{m}}{2}F_\epsilon\left(\frac{k_w-\theta}{\sigma}\right)\right) f_\epsilon\left(\frac{s-\theta}{\sigma}\right) f_\theta(\theta)d\theta}{\int_{s-\sigma\bar{\epsilon}}^{s-\sigma\underline{\epsilon}} f_\epsilon\left(\frac{s-\theta}{\sigma}\right) f_\theta(\theta)d\theta},$$

$$\pi^w(s, k_w, k_s) = \frac{\int_{s-\sigma\bar{\epsilon}}^{s-\sigma\underline{\epsilon}} \pi\left(\theta - \eta, \bar{m}F_\epsilon\left(\frac{k_w-\theta}{\sigma}\right), \frac{\bar{m}}{2}F_\epsilon\left(\frac{k_s-\theta}{\sigma}\right) + \frac{\bar{m}}{2}F_\epsilon\left(\frac{k_w-\theta}{\sigma}\right)\right) f_\epsilon\left(\frac{s-\theta}{\sigma}\right) f_\theta(\theta)d\theta}{\int_{s-\sigma\bar{\epsilon}}^{s-\sigma\underline{\epsilon}} f_\epsilon\left(\frac{s-\theta}{\sigma}\right) f_\theta(\theta)d\theta},$$

where $f_\theta(\cdot)$ is the probability density function of the prior distribution of the aggregate fundamental.

In what follows, we focus on strong banks. All derivations for weak banks are analogous. Changing the variable of integration, $z = \frac{\theta-k_s}{\sigma}$, we get

$$\pi^s(s, k_s, k_w) =$$
$$\frac{\int_{\frac{s-k_s}{\sigma}-\bar{\epsilon}}^{\frac{s-k_s}{\sigma}-\underline{\epsilon}} \pi\left(k_s + \sigma z + \eta, \bar{m}F_\epsilon(-z), \frac{\bar{m}}{2}F_\epsilon(-z) + \frac{\bar{m}}{2}F_\epsilon\left(\frac{k_w-k_s}{\sigma} - z\right)\right) f_\epsilon\left(\frac{s-k_s}{\sigma} - z\right) f_\theta(k_s + \sigma z)dz}{\int_{\frac{s-k_s}{\sigma}-\bar{\epsilon}}^{\frac{s-k_s}{\sigma}-\underline{\epsilon}} f_\epsilon\left(\frac{s-k_s}{\sigma} - z\right) f_\theta(k_s + \sigma z)dz}.$$

With infinitely small noise, we have

$$\pi^s(s, k_s, k_w) \overset{\sigma\to 0}{\to} \hat{\pi}^s(s, k_s, k_w) =$$
$$\int_{\frac{s-k_s}{\sigma}-\bar{\epsilon}}^{\frac{s-k_s}{\sigma}-\underline{\epsilon}} \pi\left(k_s + \eta, \bar{m}F_\epsilon(-z), \frac{\bar{m}}{2}F_\epsilon(-z) + \frac{\bar{m}}{2}F_\epsilon\left(\frac{k_w - k_s}{\sigma} - z\right)\right) f_\epsilon\left(\frac{s - k_s}{\sigma} - z\right) dz.$$

Denote

$$h(s, s', k_s, k_w) = \int_{\frac{s-k_s}{\sigma}-\bar{\epsilon}}^{\frac{s-k_s}{\sigma}-\underline{\epsilon}} \tilde{f}(z, s, k_s)\,\tilde{\pi}(z, s', k_s, k_w)\,dz,$$

where

$$\tilde{f}\left(z,s,k_s\right)=f_\epsilon\left(\frac{s-k_s}{\sigma}-z\right),$$

$$\tilde{\pi}\left(z,s',k_s,k_w\right)=\pi\left(s'+\eta,\bar{m}F_\epsilon\left(-z\right),\frac{\bar{m}}{2}F_\epsilon\left(-z\right)+\frac{\bar{m}}{2}F_\epsilon\left(\frac{k_w-k_s}{\sigma}-z\right)\right).$$

First, $\tilde{f}(z,s,k_s)$ satisfies the monotone likelihood ratio property by Assumption 3, namely, $\frac{\tilde{f}(z,s_H,k_s)}{\tilde{f}(z,s_L,k_s)}$ increases in z for $s_H>s_L$. Second, $\tilde{\pi}\left(z,s',k_s,k_w\right)$ satisfies the single-crossing property. That is, for a given s' there exists at most one z in which $\tilde{\pi}\left(z,s',k_s,k_w\right)$ switches sign. Therefore, by Lemma 5 in Athey (2002), $h\left(s,s',k_s,k_w\right)$ also satisfies single-crossing, that is, if $s<s^*\left(s',k_s,k_w\right)$ then $h(s,s',k_s,k_w)<0$ and if $s>s^*\left(s',k_s,k_w\right)$ then $h(s,s',k_s,k_w)>0$.

Suppose that $h(s,s,k_s,k_w)=0$. Such s exists by Assumption 2. For any $s'<s$ we have

$$h(s',s',k_s,k_w)<h(s',s,k_s,k_w)<h(s,s,k_s,k_w)=0,$$

where the first inequality holds because $h(s,s',k_s,k_w)$ increases in s' and the second inequality holds because $h(s,s',k_s,k_w)$ satisfies single-crossing. Analogously, for any $s'>s$ we have

$$h(s',s',k_s,k_w)>h(s',s,k_s,k_w)>h(s,s,k_s,k_w)=0.$$

Then there exists a cutoff $\beta_s\left(k_s,k_w\right)$ such that $\hat{\pi}^s\left(s,k_s,k_w\right)=h(s,s,k_s,k_w)$ is negative (positive) if s is below (above) the cutoff and zero at the cutoff. Similarly, $\beta_w\left(k_w,k_s\right)$ can be defined. In equilibrium, it must be that $k_s=\beta_s\left(k_s,k_w\right)$ and $k_w=\beta_w\left(k_s,k_w\right)$. Therefore, the run thresholds are implicitly given by

$$\hat{\pi}^s(\theta_s^*,\theta_s^*,\theta_w^*)=$$
$$\int_{-\bar{\epsilon}}^{-\underline{\epsilon}}\pi\left(\theta_s^*+\eta,\bar{m}F_\epsilon\left(-z\right),\frac{\bar{m}}{2}F_\epsilon\left(-z\right)+\frac{\bar{m}}{2}F_\epsilon\left(\frac{\theta_w^*-\theta_s^*}{\sigma}-z\right)\right)f_\epsilon\left(-z\right)dz=0.$$

Changing the variable of integration, $x = F_\epsilon(-z)$, we get

$$\int_0^1 \pi\left(\theta_s^* + \eta, \bar{m}x, \frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\frac{\theta_w^* - \theta_s^*}{\sigma} + F_\epsilon^{-1}(x)\right)\right) dx = 0. \tag{76}$$

Similarly, we can derive the indifference condition for marginal investors of weak banks:

$$\int_0^1 \pi\left(\theta_w^* - \eta, \bar{m}x, \frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\frac{\theta_w^* - \theta_s^*}{\sigma} + F_\epsilon^{-1}(x)\right)\right) dx = 0. \tag{77}$$

There exists a unique solution (θ_s^*, θ_w^*) to equations (76) and (77). The existence follows from Assumption 2. Below we prove the uniqueness by contradiction.

Suppose there exist two equilibria with distinct thresholds (θ_s^*, θ_w^*) and $(\hat{\theta}_s^*, \hat{\theta}_w^*)$. Without loss of generality, suppose $\hat{\theta}_s^* > \theta_s^*$. Since $\pi(z_i, m_i, m)$ increases in z_i and decreases in m, Equation (76) implies that $\frac{\hat{\theta}_w^* - \hat{\theta}_s^*}{\sigma} > \frac{\theta_w^* - \theta_s^*}{\sigma}$. Similarly, according to Equation (77), this implies $\hat{\theta}_w^* < \theta_w^*$. Since the difference between two thresholds $\hat{\theta}_w^* - \hat{\theta}_s^* > \theta_w^* - \theta_s^*$, it then must be $\hat{\theta}_s^* < \theta_s^*$, which contradicts the premise that $\hat{\theta}_s^* > \theta_s^*$. Therefore, the equilibrium is unique. $\square$

Substituting in the payoff function (75), thresholds θ_s^* and θ_w^* are determined by

$$\int_0^1 \frac{\theta_s^* + \eta - \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(\frac{\theta_w^* - \theta_s^*}{\sigma} + F_\epsilon^{-1}(x)\right)\right)\bar{m}x}{1 - \bar{m}x} dx = 1,$$
$$\int_0^1 \frac{\theta_w^* - \eta - \lambda\left(\frac{\bar{m}}{2}x + \frac{\bar{m}}{2}F_\epsilon\left(-\frac{\theta_w^* - \theta_s^*}{\sigma} + F_\epsilon^{-1}(x)\right)\right)\bar{m}x}{1 - \bar{m}x} dx = 1.$$

These are Equations (7) and (8).

Note that this proof can be straightforwardly extended to the case in which bank-specific productivity shock takes $N \geq 2$ values.