Macroeconomic Dynamics (2023), **27**, pp. 2252–2269
doi:10.1017/S1365100522000694

ARTICLE

The optimal distribution of insured and uninsured deposits across banks*

Lukas Voellmy

Swiss National Bank, Boersenstrasse 15, P.O. Box 8022, Zurich, Switzerland
Email: lukas.voellmy@snb.ch

Abstract
I study a model of self-fulfilling bank runs where government-provided deposit insurance covers small (retail) deposits but not large (wholesale) deposits. The share of the banking system that may be affected by runs depends on the distribution of retail and wholesale deposits across banks. The magnitude of runs is minimized if banks with both retail and wholesale depositors (reminiscent of commercial banks) coexist with banks that cater only to wholesale depositors (reminiscent of shadow banks). The shadow banking sector should be large enough to absorb enough wholesale deposits from commercial banks to keep them shielded from runs. In a decentralized equilibrium, the magnitude of runs tends to be larger than optimal as a result of wholesale depositors' incentive to invest in the banks with the highest share of retail depositors.

Keywords: Deposit insurance; bank runs; shadow banking

1. Introduction

Government-provided deposit insurance is an important part of financial market regulation around the world. Arguably the main reason for the prevalence of deposit insurance is its effectiveness in eliminating bank runs, which are usually considered at least partly the result of a coordination failure among depositors. Importantly for the purposes of this paper, deposit insurance coverage is limited in almost all countries with explicit deposit insurance schemes in place (Demirguc-Kunt et al. (2015)).[1] I take this feature of deposit insurance schemes as given and start from the premise that deposit insurance covers small ("retail") but not large ("wholesale") deposits.[2] The share of insured deposits among all outstanding deposits then depends on the (exogenous) share of deposits held by retail depositors.

The key contribution of this paper is to show that, for a given aggregate share of insured deposits, the potential magnitude of (systemic) bank runs depends on how insured and uninsured deposits are distributed across banks. The paper thus adds to the literature on optimal deposit insurance design by studying how insured and uninsured deposits should be distributed across banks in order to minimize the share of the banking system that can be affected by runs.

I develop a three-period model with banks that invest in an illiquid asset maturing in the final period. Banks engage in maturity transformation by issuing demand deposits that can be redeemed at par in the middle period. Banks need to liquidate assets if depositors redeem in the middle period, and the liquidation price decreases in the aggregate amount of assets sold. The

*An earlier version of this paper circulated under the title "Shadow Banking and Financial Stability under Limited Deposit Insurance". I thank Ina Bialova (discussant), Corinne Dubois (discussant), Huberto Ennis, Leonardo Gambacorta (discussant), Todd Keister, Cyril Monnet, Ewelina Laskowksa, and two anonymous referees for particularly helpful comments on various versions of the paper. The views expressed in this paper are those of the author and do not necessarily reflect those of the Swiss National Bank.

fact that early redemptions may cause fire sale losses which are then imposed on depositors who remain in the bank can lead to self-fulfilling run equilibria in the spirit of Diamond and Dybvig (1983). I abstract from fundamental risk, so that these self-fulfilling runs constitute the only source of risk in the economy.

The model features two types of depositors: "retail" depositors with a small endowment and "wholesale" depositors with a large endowment. There is an exogenous deposit insurance scheme that insures the deposits of retail but not wholesale depositors. Wholesale depositors may stand for institutional cash pools whose cash holdings far exceed the deposit insurance limit.[3] A given bank can have both retail and wholesale depositors. Since retail depositors are protected from losses, only wholesale depositors may potentially run on their bank. A bank is more likely to be susceptible to runs if a high share of its deposits is held by wholesale depositors and if the liquidation price of assets on secondary markets is low.

A key result is that in order to minimize the share of the banking system that can be affected by runs, there should only be two types of banks: (i) banks that have both retail and wholesale depositors (reminiscent of commercial banks), with the share of deposits held by retail depositors being high enough to shield these banks from runs and (ii) banks with only wholesale depositors (reminiscent of shadow banks). Importantly, the magnitude of runs in this structure of the banking system is smaller compared to an alternative situation where the share of retail and wholesale deposits is the same at all banks. If retail and wholesale deposits are distributed uniformly across banks, then systemic runs may affect the entire banking system; by concentrating wholesale deposits in one part of the banking system, runs can be limited to these banks, leaving the remaining banks with retail depositors unaffected. The optimal size of the banking sector catering only to wholesale depositors (the "shadow banking sector") is the smallest size necessary to absorb enough wholesale deposits from banks with retail depositors to keep them shielded from runs.[4]

After deriving the optimal structure of the banking system for a given aggregate share of insured deposits, I analyze the structure of the banking system that results in a decentralized equilibrium where retail and wholesale depositors choose freely in which bank to invest. Whenever systemic runs occur with positive probability, it is privately optimal for wholesale depositors to invest in the banks with the highest share of deposits held by retail depositors, since expected losses caused by runs are lowest at these banks. If the aggregate share of deposits held by wholesale depositors is high, then the optimal structure of the banking system is such that fragile banks catering only to wholesale depositors coexist with stable banks with a high share of retail deposits. This structure of the banking system does not constitute an equilibrium since it would be privately optimal for wholesale depositors to shift their deposits away from the fragile banks into the stable banks, causing the latter to become fragile as well. A simple way to implement the optimal structure of the banking system in equilibrium is to impose a ceiling on wholesale deposits held at banks that have retail depositors.

Related literature. A key result of this paper is that a heterogenous financial system – one in which banks without insured deposits coexist with banks that have a large share of insured deposits – tends to be less fragile than a homogenous financial system in which the share of insured deposits is the same at all banks. This relates the present paper to Choi (2014), Liu (2019) and Goldstein et al. (2022), all of which show how heterogeneity can foster financial stability in some sense. Compared to the papers mentioned above, which focus on heterogeneity with regard to asset returns and leverage, the mechanism studied in this paper is novel by its focus on the distribution of (inertial) insured depositors across banks.

This paper is also related to a theoretical literature studying optimal deposit insurance design, notably with regard to deposit insurance coverage. Davila and Goldstein (forthcoming) study the optimal level of deposit insurance coverage in a model where increasing coverage has the benefit of reducing expected losses caused by runs but entails social costs in the form of deadweight losses of taxation.[5] The present paper can be seen as complementary to Davila and Goldstein (forthcoming) by showing that the tradeoffs regarding the optimal level of coverage may be improved if, in addition to choosing the level of coverage, the regulator can implement an appropriate distribution of insured and uninsured deposits across banks.

Finally, this paper is related to a theoretical literature studying the financial stability effects of shadow banking.[6] Papers in this literature have emphasized different aspects of shadow banking. Plantin (2015) and Huang (2018) model shadow banks as off-balance sheet vehicles that allow commercial banks to circumvent regulatory restrictions on debt issuance, with detrimental effects for financial stability.[7] Gennaioli et al. (2013) focus on shadow banks' role in the securitization process and show how shadow banking can make the financial system more vulnerable to aggregate shocks. Gertler et al. (2016) model shadow banks as wholesale banks that can operate with higher leverage compared to retail banks due to a smaller degree of agency frictions. While shadow banking can reduce the financial accelerator in the aftermath of real shocks, high leverage in the shadow banking sector can also lead to instability in the form of bank runs. In Moreira and Savov (2017), shadow banks issue risky claims which are information-insensitive and therefore provide liquidity services. This leads to a socially desirable expansion of liquidity in normal times but makes the economy more vulnerable to changes in aggregate uncertainty. Hanson et al. (2015) and Chretien and Lyonnet (2019) characterize shadow banking and commercial banking as two different ways to provide risk-free claims. While commercial banks create risk-free claims through their access to deposit insurance, shadow banks create risk-free claims by liquidating assets immediately if bad news arrive. In Chretien and Lyonnet (2019), commercial banks act as the buyers of assets liquidated by shadow banks, which implies that an extension of deposit insurance indirectly profits shadow banks as well. Luck and Schempp (2014) develop a Diamond Dybvig-type model where commercial banks issue insured and shadow banks uninsured short-term claims. Aggregate losses caused by runs increase in the size of the shadow banking sector. The present paper is the first one to show that in a world of limited deposit insurance where traditional banks issue both insured and uninsured deposits, shadow banking might *decrease* the extent of runs and fire sales in the economy. The present paper abstracts from many issues relevant to shadow banking and should be seen as complementary to the papers mentioned above.

2. The model

Time goes on for three periods, $t = 0, 1, 2$. There is an infinitely divisible good used for consumption and investment. There are three types of agents: retail depositors, wholesale depositors, and bankers.

A measure $\sigma \in [0, 1]$ of retail depositors is born at date 0 with an endowment of one unit of good each. Additionally, a measure λ of wholesale depositors is born with an endowment of Λ units of good each, where Λ should be thought of as being much larger than 1. For convenience, depositors' aggregate endowment is normalized to one, that is, $\sigma + \lambda\Lambda = 1$, so that σ equals the share of the aggregate endowment held by retail depositors. Retail and wholesale depositors are indexed by $i_\sigma \in [0, \sigma]$ and $i_\lambda \in [0, \lambda]$, respectively. The payoff of both retail and wholesale depositors is given by $u(c_1 + c_2)$, where c_t denotes consumption at date t and $u(\cdot)$ is a continuous, strictly increasing function. The only possibility to store the endowment from date 0 to the future is through a constant-return-to-scale investment technology that returns 1 unit of good at date 2 per unit of good invested at date 0. For reasons that lie outside of the model, depositors access the investment technology only through demand deposits issued by bankers.

There is a finite number of bankers indexed by $j = 1, 2, .., \mathcal{J}$, where $\mathcal{J}$ is some large, finite number and denotes both the number of bank(er)s as well as the set of bank(er)s. Bankers are born at date 0 without endowment and they maximize $U(c_0)$, where $U(\cdot)$ is a strictly increasing function. At date 0, bankers sell demand deposits to depositors in return for depositors' endowment. For simplicity, I assume each depositor can only invest in one bank. Banks invest part of the collected

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

endowment in the investment technology and keep the remaining part as profits. I denote by $r(j)$ the gross return on deposits offered by banker j. It may be helpful to anticipate here that, as a result of Bertrand competition among banks, we will have $r(j) = 1$ for all active banks in equilibrium, so that banks make no profits.

Deposits can be redeemed at par either at date 1 or at date 2. If depositors withdraw at date 1, banks need to sell assets (claims to investment return) in order to pay out redeeming depositors. Denote by p the date 1 price of an asset paying 1 unit at date 2, and denote by μ the total fundamental value (i.e., the total date 2 payout) of all assets sold by banks at date 1. Suppose the liquidation price follows

$$p(\mu) = \begin{cases} 1 & \text{if } \mu \leqslant \overline{\mu} \\ k(\mu) & \text{if } \mu \geqslant \overline{\mu} \end{cases}, \quad \text{with} \quad 0 < \overline{\mu} < 1, \tag{1}$$

where $k(\mu)$ is continuous and satisfies $k(\overline{\mu}) = 1$, $k(\mu) > 0$ and $k'(\mu) < 0$ for $\mu \geqslant \overline{\mu}$. In words, assets can be sold at fundamental value as long as the aggregate amount of assets sold is below some threshold $\overline{\mu}$; once the amount of assets sold exceeds this threshold, the fire sale price is strictly decreasing in the amount of assets sold.[8] Both depositors and banks act as price-takers with regard to the liquidation price. For future reference, I define $p^m \equiv p(1)$ as the lowest liquidation price that may possibly occur.

The final ingredient of the model is an exogenous deposit insurance scheme that insures the deposits of retail but not wholesale depositors.[9] Banks cannot discriminate at date 0 between retail and wholesale depositors, that is, any depositor who wishes to open an account at any given bank at date 0 can do so. Whenever a bank is not able to pay out an amount of good corresponding to the face value of its retail depositors at date 2, deposit insurance steps in and repays retail depositors. Deposit insurance payments are financed by taxes which are not modeled explicitly; they are made at the end of date 2 and the deposit insurance agency remaining passive up to that point. Notably, if banks have both retail and wholesale deposits outstanding at date 2, wholesale depositors can redeem first, even if the bank incurred losses at date 1.[10]

3. Runs

In this section, I focus on depositors' decision whether or not to withdraw their deposits at date 1, taking their investment choices at date 0 as given. Since retail depositors never have an incentive to withdraw at date 1, we can limit attention to the behavior of wholesale depositors.

A bank pays out depositors who redeem at date 1 in full as long as the bank can raise enough funds to do so. If requests for redemptions exceed the liquidation value of the bank's portfolio then, following Allen and Gale (1998) and others, the bank will liquidate its entire portfolio at the going market price and distribute the proceeds equally among those who redeem. In this situation, wholesale depositors who do not redeem receive nothing. It follows that a bank will be susceptible to runs at date 1 whenever the amount of good it can raise by liquidating its entire portfolio at the going market price is insufficient to pay out all its wholesale depositors in full.

Whether a bank j is susceptible to runs depends both on the share of the bank's outstanding deposits held by retail depositors, denoted $\vartheta(j)$, as well as on the liquidation price, p. To illustrate this, consider a bank that has a measure D of outstanding deposits at date 1, which also equals the fundamental value of assets held by the bank. Suppose half the deposits are held by retail depositors, that is, $\vartheta(j) = 0.5$. Suppose further the liquidation price of assets at date 1 equals $p < 0.5$. Repaying all its wholesale depositors at date 1 would require to raise $0.5 \cdot D$ units of good at date 1. However, by liquidating its entire asset portfolio, the bank can raise only $p \cdot D < 0.5$ units of good. The bank is then susceptible to a run at date 1 since nothing will be left in the bank for wholesale depositors who do not redeem in case all (other) wholesale depositors run. Consider now a different example where, as before, 50% of the bank's deposits are held by retail depositors

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

but the liquidation price equals $p = 0.8$. The bank could then repay all its wholesale depositors at date 1 by selling a fraction $\frac{0.5}{0.8} = 0.625$ of its portfolio at the prevailing market price of $p = 0.8$. No matter how many wholesale depositors redeem at date 1, the bank has always enough funds left at date 2 to repay wholesale depositors who did not redeem at date 1. It follows that the bank is not susceptible to runs at date 1. In general, by the same reasoning, a bank j is susceptible to runs at date 1 if and only if

$$1 - \vartheta(j) > p, \tag{2}$$

That is, if and only if the share of deposits held by wholesale depositors exceeds the liquidation price.

When hit by a run, banks liquidate their entire asset portfolio which, by (1), depresses the secondary market price p. This introduces a systemic element to runs: the more banks are hit by a run, the lower p and hence the more banks are susceptible to runs. Note further that *not* redeeming deposits is always the best response if no other depositor at the same bank redeems at date 1. This implies that wholesale depositors at a given bank can always coordinate on *not* running on their bank, independent of the liquidation price p. In principle, it is thus possible that depositors coordinate on running on some subset of banks (which is large enough to cause the liquidation price p to drop below 1) and not on the remaining banks.[11] The economy may thus exhibit many different run equilibria at date 1 in which different subsets of banks are hit by a run. In the remainder of the paper, I will limit attention to two date 1 equilibria: (i) the one where nobody runs and (ii) the largest run that is possible, denoted the *systemic run*.[12] To make this precise, it is useful to introduce the following definitions:

Definition 3.1. *The **magnitude of a run** denotes the aggregate fundamental value of assets liquidated by banks at date* 1*. The **systemic run** is the run equilibrium with the largest magnitude among all run equilibria.*

It is not hard to show that the systemic run equilibrium contains all other run equilibria in the sense that the set of banks hit in the systemic run equilibrium is a strict superset of the set of banks hit in all other run equilibria. Notice that the economy may not exhibit any run equilibria, in which case the notion of systemic run is obsolete. I assume that whenever the economy does exhibit multiple date 1 equilibria, depositors select the systemic run equilibrium with probability $\pi^{sr} \in (0, 1)$ and the no-run equilibrium with probability $1 - \pi^{sr}$. Finally, it will be useful to introduce the following labels for banks with and without retail depositors, respectively:

Definition 3.2. ***Commercial banks*** *are banks with at least some retail depositors (i.e., banks with* $\vartheta(j) > 0$*).* ***Shadow banks*** *are banks without retail depositors* ($\vartheta(j) = 0$)*.*

4. The optimal structure of the banking system

In this section, I study how a social planner aiming to minimize the magnitude of systemic runs would distribute insured and uninsured deposits across banks.[13] In Section A.2. of the online appendix, I show that minimizing the magnitude of systemic runs is equivalent to maximizing the expected aggregate payout to depositors net of deposit insurance payments.

4.1. An illustrative example

We start with an example that illustrates why the magnitude of systemic runs depends on the distribution of insured (retail) and uninsured (wholesale) deposits across banks. Consider a banking system with $\sigma = 0.5$, such that 50% of all deposits are held by retail depositors. Suppose further the liquidation price (1) follows $k(\mu) = \overline{\mu}/\mu$, with $\overline{\mu} = 0.25$. Figure 1 shows three different ways how

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

Figure 1. Different distributions of retail and wholesale deposits across banks.

retail and wholesale deposits could be distributed across banks, with retail deposits being depicted in gray and wholesale deposits in white.

The left-hand side of Figure 1 shows a banking system in which retail and wholesale deposits are distributed uniformly across banks, that is, there is one representative bank with 50% retail and 50% wholesale deposits. Consider a scenario where all banks liquidate their portfolios, in which case the liquidation price falls to $p(1) = 0.25$. Since the share of wholesale depositors is higher than 25% at all banks, by condition (2), all banks are susceptible to runs at a liquidation price of $p = 0.25$. It follows that systemic runs in this economy encompass all banks.

Consider next the middle part of Figure 1, which shows a banking system with two types of banks: some with only retail depositors ("commercial banks") and some with only wholesale depositors ("shadow banks"). Consider a scenario where all shadow banks liquidate their portfolios, in which case the liquidation price falls to $p(0.5) = 0.5$. By condition (2), banks with 100% wholesale depositors are susceptible to runs at this liquidation price. The largest possible run in this banking system (the systemic run) is thus one that affects all shadow banks. Compared to the uniform distribution of deposits across banks depicted on the left-hand side of Figure 1, systemic runs encompass only half the banking system.

The extreme distribution of retail and wholesale deposits across banks depicted in the middle of Figure 1 does not in general minimize the magnitude of systemic runs. To see this, suppose that starting from the situation depicted in the middle of Figure 1, some wholesale deposits are moved from shadow banks into commercial banks, as depicted on the right-hand side of Figure 1. In the example on the right-hand side of Figure 1, the share of wholesale deposits at commercial banks equals $\frac{0.1}{0.6} < 0.25$. Since there is no scenario in which the liquidation price falls below $p(1) = 0.25$, systemic runs still do not encompass commercial banks. As before, systemic runs affect the entire shadow banking sector; however, since the size of the shadow banking sector is smaller, the magnitude of systemic runs is smaller.

4.2. Optimal distribution of deposits across banks

In this subsection, I formally derive the structure of the banking system that minimizes the magnitude of systemic runs. To start, let $D(\mathcal{J}')$ denote the total face value of deposits held at some subset of banks $\mathcal{J}' \subseteq \mathcal{J}$, which is the same as the fundamental value of assets held by these banks. For each bank, the planner chooses the total amount of deposits held at the bank, $D(j)$, as well as the share of retail deposits at the bank, $\vartheta(j)$.

In a scenario where all banks in $\mathcal{J}'$ are hit by a run, they all liquidate their assets, so that the liquidation price falls to $p(D(\mathcal{J}')) \leqslant 1$. From the discussion in Section 3, it follows that the economy exhibits a run equilibrium encompassing all banks in $\mathcal{J}'$ iff the share of wholesale deposits at each bank in $\mathcal{J}'$ is higher than the liquidation price in a scenario where all banks in $\mathcal{J}'$ liquidate their portfolios. The condition for a run equilibrium encompassing all banks $j \in \mathcal{J}'$ to exist is thus

$$1 - \vartheta(j) > p(D(\mathcal{J}')) \qquad \text{for all banks} \quad j \in \mathcal{J}'. \tag{3}$$

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

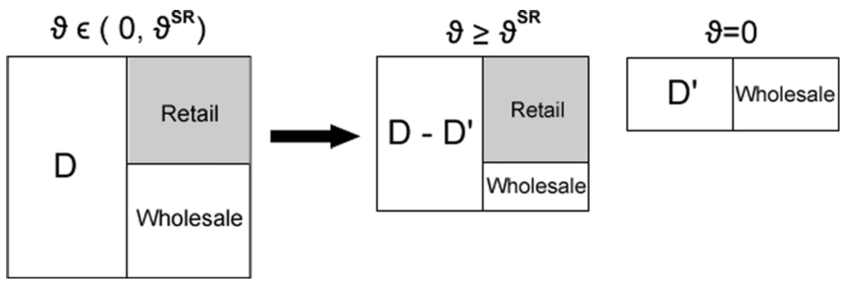

Figure 2. Illustration of the proof of Proposition 4.1.

The following result shows that the systemic run encompasses all banks whose share of retail deposits is below some threshold ϑ^{sr}:

Lemma 4.1. *For any given structure of the banking system* $\{D(j), \vartheta(j)\}_{j\in\mathcal{J}}$, *there exists a threshold* $\vartheta^{sr} \in [0, 1]$ *such that a systemic run encompasses all banks* $j \in \mathcal{J}$ *with* $\vartheta(j) < \vartheta^{sr}$ *and does not encompass banks with* $\vartheta(j) \geqslant \vartheta^{sr}$.

The proof of Lemma 4.1 is an application of Tarski's fixed point theorem and is given in Appendix A.1. Using the result of Lemma 4.1, we can formulate the planner's problem as:

$$\begin{aligned} &\min_{\{D(j),\vartheta(j)\}_{j\in\mathcal{J}}} \quad D\left(\{j \in \mathcal{J} \mid \vartheta(j) < \vartheta^{sr}\}\right) \\ &\text{s.t.} \sum_{j\in\mathcal{J}} D(j) = 1 \quad \text{and} \quad \sum_{j\in\mathcal{J}} \vartheta(j)D(j) = \sigma \end{aligned} \tag{4}$$

We then get the following result:

Proposition 4.1. *Minimizing the magnitude of systemic runs requires* $D(\{j \in \mathcal{J} \mid \vartheta(j) = 0\}) + D(\{j \in \mathcal{J} \mid \vartheta(j) \geqslant \vartheta^{sr}\}) = 1$.

Stated verbally, Proposition 4.1 says that minimizing the magnitude of system runs requires that there be only two types of banks: banks without retail deposits and banks with a high enough share of retail deposits to prevent them from being susceptible to runs. The formal proof of Proposition 4.1 is given in Appendix A.2. Figure 2 provides a graphical illustration of the proof. Suppose there are banks with $\vartheta \in (0, \vartheta^{sr})$ that issue some amount $D > 0$ of deposits. The fundamental value of assets liquidated by these banks in a systemic run equals D. Suppose we reallocate an amount D' of wholesale deposits away from these banks into newly created banks, where D' is such that the share of retail deposits at the original banks reaches ϑ^{sr}. An amount $D - D'$ of deposits is now held at banks that are run-proof, which decreases the amount of assets liquidated in a systemic run. Therefore, the initial structure of the banking system cannot minimize the magnitude of systemic runs.

In the following, without loss of generality, I assume the share of retail deposits at all commercial banks is set to the same level ϑ_{CB}.[14] From Proposition 4.1, we know that minimizing the magnitude of systemic runs requires $\vartheta_{CB} \geqslant \vartheta^{sr}$. We continue with the following result, whose proof is given in Appendix A.3:

Lemma 4.2. *Suppose* $\vartheta(j) \in \{0, \vartheta_{CB}\}$ *for all banks* $j \in \mathcal{J}$. *Then* $\vartheta_{CB} \geqslant \vartheta^{sr}$ *is equivalent to* $\vartheta_{CB} \geqslant 1 - p^m$.

To streamline notation, denote $D_{CB} \equiv D(\{j \in \mathcal{J} \mid \vartheta(j) > 0\})$ as the (relative) size of the commercial banking sector and $D_{SB} \equiv D(\{j \in \mathcal{J} \mid \vartheta(j) = 0\})$ as the (relative) size of the shadow banking sector. Since the aggregate share of deposits held by retail depositors equals σ, we

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

have $D_{CB}\,\vartheta_{CB} = \sigma$. Combining this with $\vartheta_{CB} \geqslant 1 - p^m$ (from Lemma 4.2) and substituting $D_{SB} = 1 - D_{CB}$ yields the following condition on the size of the shadow banking sector:

$$D_{SB} \geqslant \max\left\{1 - \frac{\sigma}{1 - p^m}, 0\right\} \equiv D_{SB}^{min}(\sigma) \tag{5}$$

In words, $D_{SB}^{min}(\sigma)$ is the minimum size of the shadow banking sector required to absorb enough wholesale deposits from the commercial banking sector such as to keep commercial banks shielded from runs. Notice that $D_{SB}^{min}(\sigma)$ is decreasing in σ; if σ is high enough, we have $D_{SB}^{min}(\sigma) = 0$, that is, commercial banks are not susceptible to runs even if all wholesale deposits are held at commercial banks.

So far we have seen that if σ is relatively low, then shadow banks need to have some minimum size in order to keep commercial banks stable. But at which point do shadow banks themselves become susceptible to runs? By condition (2), shadow banks are susceptible to runs whenever the liquidation price p drops below one. Given that commercial banks are never hit by a run, this implies that shadow banks are susceptible to systemic runs iff the liquidation price drops below one in a situation where all shadow banks are hit by a run, which, by (1), is the case iff $D_{SB} > \overline{\mu}$. We can therefore denote $D_{SB}^{max} \equiv \min\{1 - \sigma, \overline{\mu}\}$ as the maximum size of the shadow banking sector for which shadow banks are not susceptible to systemic runs, given that commercial banks are not susceptible to runs.[15] The discussion in the previous paragraphs leads us to the following proposition, whose proof is given in Appendix A.4:

Proposition 4.2. *Denote D_{SB}^{opt} as the (relative) size of the shadow banking sector for which the magnitude of systemic runs is minimized. D_{SB}^{opt} depends on σ as follows:*

(i) If $\sigma \in [(1 - p^m), 1]$, then $D_{SB}^{opt} \in [0, D_{SB}^{max}]$.

(ii) If $\sigma \in [(1 - \overline{\mu})(1 - p^m), (1 - p^m)]$, then $D_{SB}^{opt} \in [D_{SB}^{min}(\sigma),\ D_{SB}^{max}]$.

(iii) If $\sigma \in [0, (1 - \overline{\mu})(1 - p^m)]$, then $D_{SB}^{opt} = D_{SB}^{min}(\sigma)$.

According to Proposition 4.2, σ can be divided into three regions. In region (i), we get from condition (5) that $D_{SB}^{min}(\sigma) = 0$, meaning that commercial banks are not susceptible to systemic runs even if all wholesale deposits are held in the commercial banking sector. Systemic runs then do not occur in the economy as long as the size of the shadow banking sector is below D_{SB}^{max}. Next, in region (ii), we have $D_{SB}^{min}(\sigma) > 0$, that is, the relative size of the shadow banking sector must be larger than zero in order to absorb enough wholesale deposits from the commercial banking sector to keep commercial banks stable. Since $D_{SB}^{min}(\sigma) \leqslant D_{SB}^{max}$, systemic runs can be avoided by setting the relative size of the shadow banking sector within $[D_{SB}^{min}(\sigma), D_{SB}^{max}]$. Finally, in region (iii), we have $D_{SB}^{min}(\sigma) > D_{SB}^{max}$, that is, the minimum size of the shadow banking sector required to keep commercial banks stable is such that shadow banks are susceptible to systemic runs. Preventing runs entirely is then not feasible, and the magnitude of systemic runs is minimized by setting the shadow banking sector to the smallest size necessary to keep commercial banks stable.

We can also derive from Proposition 4.2 what fraction of the banking system can be shielded from systemic runs for given σ. For $\sigma \geqslant (1 - \overline{\mu})(1 - p^m)$, systemic runs can be prevented entirely by distributing retail and wholesale deposits optimally across banks. If $\sigma < (1 - \overline{\mu})(1 - p^m)$, the maximum fraction of the banking system that can be kept run-proof corresponds to the relative size of the (stable) commercial banking sector in the optimal structure of the banking system, which equals $1 - \sigma_{SB}^{min}(\sigma) = \dfrac{\sigma}{1 - p^m}$. This leads to the following result:

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

Corollary 4.1. *Denote $S(\sigma)$ as the maximum share of the banking system that can be shielded from systemic runs for given σ. Then:*

$$S(\sigma) = \begin{cases} 1 & \text{for } \sigma \in \left[(1-\overline{\mu})(1-p^m), 1\right] \\ \dfrac{\sigma}{1-p^m} & \text{for } \sigma \in \left[0, (1-\overline{\mu})(1-p^m)\right) \end{cases}$$

5. Date-0 equilibrium

In this section, I examine what structure of the banking system results in a decentralized equilibrium at date 0 where banks offer deposit contracts and depositors choose in which bank to invest. I then compare the structure of the banking system resulting in a decentralized date 0 equilibrium to the optimal structure derived in Section 4.

In a decentralized allocation, each bank j sets the gross return $r(j) \in [0, 1]$ of the deposit contract it offers. The bank chosen by a given retail depositor i_σ is denoted by $\alpha_\sigma(i_\sigma) \in \mathcal{J}$, and $\alpha_\lambda(i_\lambda) \in \mathcal{J}$ denotes the same for a given wholesale depositor i_λ. The equilibrium definition is standard: A date-0 equilibrium is a set of deposit contracts $\{r(j)\}_{j\in\mathcal{J}}$ together with a set of investment choices $\{\alpha_\sigma(i_\sigma)\}_{i_\sigma\in[0,\sigma]}$ and $\{\alpha_\lambda(i_\lambda)\}_{i_\lambda\in[0,\lambda]}$ such that (i) each depositor chooses the bank that gives her the highest expected payoff and (ii) no bank can strictly increase profits by changing the return paid on deposits.

Equilibrium behavior of banks and retail depositors is straightforward. Since retail depositors are guaranteed to receive the offered return $r(j)$, they always invest in the bank offering the highest return, and they are indifferent between all banks that offer the same return. Furthermore, it is not hard to see that, by the usual argument of Bertrand, all banks that are active in equilibrium offer a return $r(j) = 1$, that is, banks pay a return equal to the return of the investment technology and make zero profits in equilibrium. To streamline the exposition, I will take this as given from now on.

What makes the decentralized equilibrium interesting are the investment choices of wholesale depositors. In order to write down expected payoffs of wholesale depositors, we need to introduce some additional notation. First, we denote the liquidation price in a systemic run by

$$p^{sr} \equiv p\left[D\left(\{j \in \mathcal{J} \mid \vartheta(j) < \vartheta^{sr}\}\right)\right]. \tag{6}$$

Next, we denote by $\varphi(\vartheta, p)$ the fraction of wholesale deposits (in terms of their face value) a bank can pay out when hit by a run. The share of wholesale deposits a bank can repay in a run equals the amount good the bank can raise by liquidating its entire asset portfolio at price p, divided by the face value of wholesale deposits held at the bank. We thus have

$$\varphi(\vartheta, p) = \min\left\{\frac{p}{1-\vartheta}, 1\right\}. \tag{7}$$

When choosing which bank to invest in, an individual wholesale depositor takes the distribution of retail and wholesale deposits across banks as well as the liquidation price in a systemic run as given. Denote $\tilde{r}(j)$ as the effective return on wholesale deposits at bank j. The no-run equilibrium is selected with probability $1-\pi^{sr}$, in which case wholesale deposits at all banks yield a gross return of one. With probability π^{sr}, a systemic run occurs, in which case the effective gross return on wholesale deposits held at bank j equals $\varphi(\vartheta(j), p^{sr})$. We thus have

$$\tilde{r}(j) = (1-\pi^{sr})u(1) + \pi^{sr}u\left(\varphi(\vartheta(j), p^{sr})\right). \tag{8}$$

If systemic runs do not occur in the economy, then we have $\varphi(\cdot) = 1$ for all banks, which implies $\tilde{r}(j) = 1$ for all banks. If systemic runs do occur, then $\tilde{r}(j)$ is strictly higher for banks with a higher share of retail deposits, which means that wholesale depositors will invest in the banks

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

with the highest share of retail deposits. Before we proceed, it is useful to introduce the following definitions:

Definition 5.1. *A date-0 equilibrium is said to be of* ***type A*** *if $\vartheta(j) \geqslant \vartheta^{sr}$ for all banks and is said to be of* ***type B*** *if $\vartheta(j) < \vartheta^{sr}$ for all banks.*

In words, a type A equilibrium is an equilibrium in which systemic runs do not occur, and a type B equilibrium is an equilibrium in which systemic runs affect all banks. We continue with the following result:

Proposition 5.1. *Every date-0 equilibrium is either of type A or of type B. In a type B equilibrium, the share of retail deposits equals σ at each bank.*

The proof of Proposition 5.1 is relatively straightforward, which is why a formal proof is omitted. First, note that if the equilibrium is such that systemic runs do occur, then systemic runs must affect all banks. Otherwise, wholesale depositors would move away from the banks that are susceptible to runs to those that are not. Further, if all banks are susceptible to systemic runs, the share of retail deposits must be the same across all banks. Otherwise, wholesale depositors would move from the banks with a lower share of retail deposits to those with a higher share. Finally, if all banks have the same share of retail deposits, the share of retail deposits at any given bank must equal the aggregate share σ.

We continue with the following result, whose proof is given in Appendix A.5:

Proposition 5.2. *The set of date-0 equilibria depends on σ as follows:*

(i) For $\sigma \in [0, (1-\overline{\mu})(1-p^m))$ only type B equilibria exist.
(ii) For $\sigma \in [(1-\overline{\mu})(1-p^m), (1-p^m))$ both type A and type B equilibria exist.
(iii) For $\sigma \in [(1-p^m), 1]$ only type A equilibria exist.

Consider first item (i) in Proposition 5.2. We know from Section 4 that preventing runs entirely is not feasible in this region of σ and that runs affect only part of the banking system if deposits are distributed optimally. In a decentralized date-0 equilibrium, runs affect the entire banking system; the optimal structure of the banking system in which fragile shadow banks coexist with stable commercial banks does not constitute a date 0 equilibrium since wholesale depositors have a private incentive to move from fragile shadow banks into stable commercial banks, causing the latter to become fragile as well.

Consider next region (ii) in which multiple date 0 equilibria exist. As seen in Section 4, systemic runs can be avoided in this region of σ if the shadow banking sector has the right size. In a situation where neither commercial nor shadow banks are susceptible to runs, wholesale depositors are indifferent between the two, so that this constitutes a date 0 equilibrium. However, once we get to a situation where systemic runs do occur, wholesale depositors will flock to the banks with the highest shares of retail deposits, and we end up in a date 0 equilibrium where all banks have the same share of retail deposits. Note that in the "good" type A equilibrium in region (ii), the fact that commercial banks are not susceptible to runs sustains the liquidation price of assets. This in turn implies that shadow banks are not susceptible to runs either, since they can sell assets on the secondary market without incurring fire sale losses.[16] Lastly, item (iii) in Proposition 5.2 represents the case where commercial banks are not susceptible to runs even if all wholesale deposits are held at commercial banks and need no further elaboration.

Even though the transition from one equilibrium to another is not explicitly modeled, we can discuss informally the stability of the type A equilibrium in region (ii). According to Proposition 4.2, systemic runs do not occur in this region of σ iff the relative size of the shadow banking sector is within $D_{SB} \in [D_{SB}^{min}(\sigma), D_{SB}^{max}]$. This implies that for any $D_{SB} \in (D_{SB}^{min}(\sigma), D_{SB}^{max})$, the type A equilibrium is stable in the sense that small changes in the relative size of the shadow

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

banking sector will not cause the economy to move to a type B equilibrium. If either $D_{SB} > D_{SB}^{max}$ or $D_{SB} < D_{SB}^{min}(\sigma)$, then systemic runs occur. In the first case ($D_{SB} > D_{SB}^{max}$) systemic runs only affect shadow banks. Depositors would then move away from fragile shadow banks into stable commercial banks, which would cause D_{SB} to fall below D_{SB}^{max} again, so that the economy would move back to the type A equilibrium. In the second case ($D_{SB} < D_{SB}^{min}$) both commercial and shadow banks are fragile. Since losses caused by runs are smaller at commercial banks, depositors would move away from shadow into commercial banks, causing the economy to transition to a type B equilibrium in which only fragile commercial banks exist.

5.1. Implementing the optimal distribution of deposits across banks

This subsection discusses how the optimal structure of the banking system could be implemented in a date 0 equilibrium, for any value of σ. A simple intervention to achieve this is a ceiling on the share of wholesale deposits held at commercial banks:

Proposition 5.3. *Suppose there is a ceiling on the share of wholesale deposits held at commercial banks, such that every bank with $\vartheta(j) > 0$ must have $\vartheta(j) \geqslant \underline{\vartheta} \equiv 1 - p^m$. Then every date-0 equilibrium (with the ceiling as an additional constraint) implements the optimal structure of the banking system.*

To proof Proposition 5.3, note first that the ceiling is such that commercial banks are not susceptible to runs even if the liquidation price drops to its lower bound p^m. It follows immediately that the ceiling is sufficient to eliminate type B equilibria. From the result in Lemma 4.2, it follows that the ceiling is also necessary to eliminate type B equilibria whenever $\sigma < 1 - p^m$. Since $D_{CB}\vartheta_{CB} = \sigma$, the size of the commercial banking sector under the ceiling satisfies $D_{CB} \leqslant \sigma/\underline{\vartheta}$ whenever $\sigma < 1 - p^m$, such that the size of the shadow banking sector satisfies $D_{SB} \geqslant 1 - (\sigma/\underline{\vartheta}) \equiv D_{SB}^{min}(\sigma)$. For $\sigma \in [(1 - \overline{\mu})(1 - p^m), (1 - p^m))$, the ceiling eliminates type B equilibria and we are left with only type A equilibria. For $\sigma < (1 - \overline{\mu})(1 - p^m)$, we have $D_{SB}^{min}(\sigma) > D_{SB}^{max}$, which means shadow banks are prone to runs under the ceiling. The ceiling then binds in the sense that more of the wholesale depositors would like to hold deposits at commercial banks but are prevented from doing so by the ceiling. In this case, the equilibrium size of the shadow banking sector equals the smallest size consistent with the ceiling, $D_{SB}^{min}(\sigma)$, which is also the optimal size.

The result of Proposition 5.3 is reminiscent of Farhi et al. (2009), where optimal regulation takes the form of a ceiling on illiquid assets (or a floor on liquid assets), that is, an upper bound on the portfolio share held in the illiquid asset. One difference is that the regulation in Farhi et al. (2009) is on the asset side rather than on the liability side as in this paper. In Farhi et al. (2009), the ceiling on illiquid assets increases the date 1 market price of illiquid assets, which allows to provide optimal insurance against liquidity shocks in the presence of unobservable side trades. In the present paper, the ceiling on uninsured deposits at commercial banks ensures that at least some banks are shielded from runs, which also sustains the date 1 price of illiquid assets. The two regulatory interventions could be seen as complementary since they address different inefficiencies. Note however that regulation in the present paper is necessarily on the liability side given that there is only one (illiquid) asset. If banks could choose between a liquid and an illiquid asset, a ceiling on illiquid assets may have a similar effect as the upper bound on uninsured deposits. It keeps banks subject to the ceiling shielded from runs and thereby sustains the date 1 price of illiquid assets.

6. Discussion of model assumptions

6.1. The role of fire-sales

I assumed the liquidation price decreases in the aggregate amount of assets sold. In order to get some insight into the role of fire sales in the model, consider an alternative set-up where each

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

bank invests in a technology that can be liquidated for some fixed amount of good at date 1, say $q \in (0, 1)$ units per unit invested at date 0. In this alternative set-up, runs at one bank do not create spill-over effects on other banks. Instead, following the discussion in Section 3, each bank j will be susceptible to runs iff $\vartheta(j) < q$, independent of what happens at other banks.

It is not hard to see that some key results regarding the optimal structure of the banking system would continue to hold in the set-up without fire sales. Consider first the case $\sigma < q$. If deposits are distributed uniformly across banks, all banks will be susceptible to runs. In order to minimize the amount of assets liquidated in a run, there should be two types of banks: banks with $\vartheta(j) = q$ and banks with $\vartheta(j) = 0$. The total share of deposits held at banks with $\vartheta(j) = 0$ should be the minimum necessary to maintain $\vartheta(j) = q$ at the other banks. Different to the setting with fire sales, banks with $\vartheta(j) = 0$ will always be prone to runs, and eliminating runs entirely is not feasible for $\sigma < q$. The reason is that, without fire sales, keeping some (commercial-) banks stable does not have the added benefit of keeping the remaining (shadow-) banks stable as well. Consider next the case with $\sigma \geqslant q$. In this case, runs can be prevented entirely by keeping the share of wholesale deposits at each individual bank below q. Different to the case with fire sales, a more homogenous distribution of deposits across banks then strictly dominates a more heterogeneous distribution.

Some key features of the competitive equilibria studied in Section 5 would continue to hold as well in the case without fire sales. In particular, wholesale depositors would still have an incentive to invest in the banks with the highest share of retail deposits due to lower losses caused by runs at these banks. This implies that for $\sigma \geqslant q$, date 0 equilibria will be of type A, while for $\sigma < q$, they will be of type B. Different to the case with fire sales, there will be no region of σ with both type A and B equilibria. The reason is again that spill-over effects are not present, and strategic complementarities exist only at the level of individual banks, not between banks.

6.2. Seniority of insured deposits

I assumed wholesale depositors can redeem before retail depositors at date 2 even if a bank is insolvent. If instead the regulator seized the bank at date 2 and treated uninsured deposits as junior to insured deposits (as the FDIC is mandated to do), then any liquidation losses the bank incurs at date 1 would be imposed first on wholesale depositors who still hold a deposit in the bank at date 2. Wholesale depositors then would have an incentive to withdraw at date 1 whenever they expect the bank to incur liquidation losses, however small. Hence, each bank with $\vartheta(j) < 1$ would be susceptible to runs whenever $p < 1$; the exact share of wholesale deposits held at the bank would be irrelevant.[17]

Assuming that retail deposits constitute a senior claim to the bank's assets at date 2 would make wholesale depositors much more prone to running at date 1 but would otherwise not fundamentally change the results. The extent of runs would be minimized by having (commercial-) banks with only retail depositors ($\vartheta(j) = 1$) coexist with (shadow-) banks with only wholesale depositors. Preventing runs completely would be feasible iff $1 - \sigma \leqslant \overline{\mu}$, that is, iff the liquidation price does not fall below one if all shadow banks liquidate their assets. In a decentralized equilibrium, wholesale depositors would still have an incentive to invest in the banks with the highest share of retail deposits, and the economy may end up in a suboptimal type B equilibrium. The optimal structure of the banking system could be implemented in equilibrium by banning completely the issuance of wholesale deposits by commercial banks.

A condition similar to (2) could still be obtained in a set-up where retail deposits are senior to wholesale deposits at date 2, if one assumes additionally that banks have some buffer to absorb losses. A more general condition capturing the essence of Condition (2) would be to say that a bank is susceptible to runs whenever $F(\vartheta(j), p) < 0$, where $F(\cdot)$ is continuous and strictly increasing in both arguments, with $F(1, 1) > 0$. Suppose that, besides deposits, banks issue equity at date 0 that is junior to wholesale deposits. Wholesale depositors would then have an incentive to run iff the liquidation losses incurred by the bank at date 1 exceed the outstanding equity. Since liquidation losses are continuously decreasing in the liquidation price and continuously increasing in the amount of wholesale deposits withdrawn, a condition as above would result.

6.3. Discriminating between retail and wholesale depositors

Another simplifying assumption made in the paper is that each bank can only post one deposit contract and all depositors, both retail and wholesale, can open an account at the bank. This raises the question what would happen if banks could offer different deposit returns for the two types of depositors or stop accepting depositors of a given type.

Fully working out the competitive equilibrium for the case where banks can discriminate between retail and wholesale depositors would require a significant change in the set-up and is beyond the scope of this paper. It is clear that type A equilibria would continue to exist as before; if systemic runs do not occur, all deposits are risk-free and competition among banks will lead banks to offer a return of one on all deposits. Things are more complicated in equilibria where runs occur. In particular, wholesale depositors would then be willing to accept a lower promised return $r(j)$ at banks with lower expected losses caused by runs. This opens up the possibility that banks compete for retail depositors (e.g., by offering them a gross return strictly higher than one) since the presence of retail depositors allows banks to lower the promised return on wholesale deposits. How exactly this plays out likely depends on some additional assumptions one would need to make, for example, about the precise way how banks compete with each other, whether there are regulatory restrictions on the return offered on insured deposits, and so on.

6.4. Deposit freezes and other mechanisms to prevent runs

In order to focus on the themes of most interest for this paper, I have taken it as given that banks issue demand deposits (as is observed in the real world) rather than deriving an optimal contract endogenously. Within the model, banks could eliminate runs at no cost by simply not allowing any redemptions at date 1 since no depositor truly needs to consume at date 1. If there were some "impatient" depositors with consumption needs at date 1, and the share of impatient depositors was known ex ante, then runs could be eliminated by freezing deposits after a certain amount of redemptions (Diamond and Dybvig (1983)). However, even when maintaining the assumption of no aggregate risk, there are a number of reasons why deposit freezes may not prevent runs: freezes may lead to preemptive runs (Engineer (1989)), they may not be credible (Ennis and Keister (2009)) or they may create losses on bank assets by lowering demand on goods markets (Altermatt et al. (2022)). To the extent that imposing a freeze is costly for an individual bank, there may also be a coordination problem: if all banks prevent runs with freezes, assets can be sold without loss at date 1, in which case it is individually optimal for banks not to impose a freeze. Nevertheless, one should keep in mind that banks may in principle be able to eliminate run equilibria without the help of a public safety net by using more sophisticated mechanisms, for instance by incentivizing depositors to signal when they are running (Andolfatto et al. (2017), Cavalcanti and Monteira (2016)) or by charging carefully calibrated redemption fees (Voellmy (2021)).

7. Conclusion

This paper presents a theoretical argument why the distribution of insured (retail) and uninsured (wholesale) deposits across banks matters for financial stability. More specifically, the paper shows that if a large part of deposits is held by wholesale cash investors without access to deposit insurance, then the presence of financial institutions that cater exclusively to these investors can be beneficial from a financial stability point of view. Concentrating uninsured short-term claims in one part of the financial system ensures that uninsured short-term debt does not destabilize the parts of the financial system that cater to insured retail depositors as well.

The paper also adds to the debate on the effects of shadow banking on financial stability. The fact that shadow banks, such as money market mutual funds in the USA, issue "money-like" liabilities without deposit insurance protection is often said to contribute to financial instability. What is sometimes missed in this debate is the fact that shadow banks cater largely to institutional cash pools managing cash balances that far exceed the deposit insurance cap (Poszar (2011), Claessens et al. (2012)). A large part of the funds invested in shadow banks would thus likely not be covered by deposit insurance even if invested in commercial banks. The results in this paper suggest that the effect of shadow banking on financial stability is ambiguous if deposit insurance coverage is limited. Given that a large share of deposits is effectively "uninsurable", the counterfactual situation without shadow banks and a much larger amount of uninsured deposits held at commercial banks may not necessarily be preferable from a financial stability perspective.

Notes

1 The limit on coverage takes various forms, such as a cap on insured deposits per depositor and bank (the most common type), a cap per depositor, or a cap per account.

2 The reasons why a blanket insurance of all deposits is usually not considered optimal are outside the scope of this paper. Among the most commonly named downsides of deposit insurance are the creation of moral hazard and the potential fiscal costs of a credible guarantee of all insured deposits.

3 Poszar (2011) discusses the fact that a significant portion of institutional cash in the USA is effectively "uninsurable" due to the deposit insurance cap.

4 This result follows the well-known principle of "narrow victories, landslide losses", which also features prominently in the literature on optimal partisan gerrymandering. In brief, a partisan gerrymanderer aiming to maximize the number of districts he wins creates exactly two types of districts: districts he wins with the smallest possible margin and districts he wins with overwhelming margin. See Kolotilin and Wolitzky (2020) for an overview. I thank an anonymous referee for pointing out the similarity between this paper's findings and the solution to the optimal gerrymandering problem.

5 Dreyfus et al. (1994) and Manz (2009) also study the optimal level of deposit insurance coverage. The basic tradeoffs are similar to Davila and Goldstein (forthcoming). Manz (2009) puts more emphasis on the moral hazard effects of increasing deposit insurance coverage while Dreyfus et al. (1994) emphasize that the optimal level of coverage depends on other aspects of public policy such as implicit bail-out guarantees and regulatory forbearance.

6 In this paper, shadow banks simply denote banks with no insured depositors, while banks with at least some insured depositors are called commercial banks. This is a highly stylized representation of the difference between commercial and shadow banks since, in reality, the two types of financial institutions differ in more dimensions than just whether or not their liabilities are (explicitly) insured by the government. Nevertheless, the paper highlights potentially important aspects of shadow banking in the context of limited deposit insurance that have not been analyzed so far.

7 This "regulatory arbitrage" view on shadow banking is also taken by a number of papers that focus on the impact of unregulated shadow banking on optimal regulation of (commercial) banks (Grochulski and Zhang (2019), Clayton and Schaab (2022), Farhi and Tirole (2021)).

8 In section A.1 of the online appendix, I show that a price function satisfying (1) with $k(\mu) = \overline{\mu}/\mu$ results endogenously if one assumes that competitive outside investors with an aggregate endowment of $\overline{\mu}$ buy assets at date 1 ("cash-in-the-market pricing").

9 I abstract from the fact that wholesale depositors may obtain deposit insurance for a significant fraction of their cash holdings by spreading deposits over many – possibly hundreds of – banks. More generally, modeling the deposit insurance cap as a cap per depositor and bank would call for a richer model in which the number of banks is itself an equilibrium outcome.

10 See Subsection 6.2 for a discussion of this assumption. I abstract from any other issues related to moral hazard. In particular, bankers cannot divert funds at date 0, that is, they must undertake an amount of real investment that corresponds to the face value of the deposits they issue.

11 Strategic complementarities with regard to withdrawal decisions exist both *within* and *across* banks. Bank runs models that feature both within- and cross-bank complementarities in a similar manner include Liu (2016), Liu (2019), Goldstein et al. (2022) and Li and Ma (2022).

12 One may imagine that equilibrium selection at date 1 is driven by a binary sunspot variable. It is then natural to assume that either nobody runs or everybody runs.

13 I take it as given in the planner's problem that depositors' entire endowment is invested in banks at date 0, so that the aggregate face value of outstanding deposits equals 1. This reflects the fact that the present paper is not concerned with the optimal amount of total short-term claims but rather with the optimal distribution of a given amount of insured and uninsured short-term claims across banks.

14 Assuming an identical share of retail deposits at all commercial banks is without loss of generality in the following sense. Consider any distribution of deposits across banks that satisfies the condition in Proposition 4.1. Keeping the total shares of investment allocated to commercial and shadow banks fixed, the magnitude of systemic runs is minimized if all commercial banks have the same share of insured deposits. To see this, note that commercial banks are not susceptible to runs as long as the distribution of deposits across banks satisfies the condition in Proposition 4.1. The distribution of retail and wholesale deposits across commercial banks therefore does not affect the magnitude of systemic runs, so that assuming a uniform distribution of deposits across commercial banks is without loss of generality.

15 Note that the size of the shadow banking sector cannot exceed $1 - \sigma$ by definition.

16 This is reminiscent of Choi (2014) who develops a model in which government support for the stronger financial institutions is more effective in maintaining financial stability than support for weaker institutions. The basic intuition in Choi (2014) is similar to this paper: supporting the strongest financial institutions sustains the liquidation price on the secondary market, which makes weaker financial institutions less susceptible to runs as well. A similar result is also found by Goldstein et al. (2022).

17 One might add here that seizing every insolvent bank's assets at date 2 and not allowing any further payouts to wholesale depositors may not be in the best interest of the regulator since it maximizes wholesale depositors' ex ante incentive to run for a given liquidation price p. In contrast, the laissez-faire policy assumed in this paper minimizes wholesale depositors' incentives to run for given p but maximizes losses to the deposit insurance fund given that a run occurs. Schilling (2018) studies a much more elaborate version of a similar tradeoff and derives a resolution authority's optimal forbearance policy for a given deposit insurance cap.

References

Allen, F. and D. Gale (1998) Optimal financial crises. *Journal of Finance* 53(4), 1245–1284.

Altermatt, L., H. van Buggenum and L. Voellmy (2022) *Systemic bank runs without aggregate risk: how a misallocation of liquidity may trigger a solvency crisis*, SNB Working Paper, 10/2022.

Andolfatto, D., E. Nosal and B. Sultanum (2017) Preventing bank runs. *Theoretical Economics*, 12, 1003–1028.

Cavalcanti, R. and P. K. Monteira (2016) Enriching information to prevent bank runs. *Economic Theory* 62(3), 477–494.

Choi, D. B. (2014) Heterogeneity and stability: Bolster the strong, not the weak. *The Review of Financial Studies* 27(6), 1830–1867.

Chretien, E. and V. Lyonnet (2019) *Are Traditional and Shadow Banks Symbiotic?*Fisher College of Business Working Paper No, 2019-03-011.

Claessens, S., Z. Poszar, L. Ratnovski and M. Singh (2012) Shadow banking: Economics and policy. *IMF Staff Discussion Note* No. 12/12.

Clayton, C. and A. Schaab (2022) *Regulation with Externalities and Misallocation in General Equilibrium*, Working Paper.

Davila, E. and I. Goldstein (forthcoming) Optimal deposit insurance. *Journal of Political Economy*.

Demirguc-Kunt, A., E. Kane and L. Laeven (2015) Deposit insurance around the world: A comprehensive analysis and database. *Journal of Financial Stability* 20, 155–183.

Diamond, D. W. and P. H. Dybvig (1983) Bank runs, deposit insurance and liquidity. *Journal of Political Economy* 91(3), 401–419.

Dreyfus, J.-F., A. Saunders and L. Allen (1994) Deposit insurance and regulatory forbearance: Are caps on insured deposits optimal? *Journal of Money, Credit and Banking* 26(3), 412–438.

Engineer, M. (1989) Bank runs and the suspension of deposit convertibility. *Journal of Monetary Economics* 24(3), 443–454.

Ennis, H. and T. Keister (2009) Bank runs and institutions: The perils of intervention. *American Economic Review* 99(4), 1588–1607.

Farhi, E., M. Golosov and A. Tsyvinski (2009) A theory of liquidity and regulation of financial intermediation. *Review of Economic Studies* 76(3), 973–992.

Farhi, E. and J. Tirole (2021) Shadow banking and the four pillars of traditional financial intermediation. *The Review of Economic Studies* 88(6), 2622–2653.

Gennaioli, N., A. Shleifer and R. W. Vishny (2013) A model of shadow banking. *The Journal of Finance* 68(4), 1331–1363.

Gertler, M., N. Kiyotaki and A. Prestipino (2016) Wholesale banking and bank runs in macroeconomic modelling of financial crises. In: Gertler, M., N. Kiyotaki and A. Prestipino.(eds.), *Handbook of Macroeconomics*, vol. 2B, pp. 1345–1425. Amsterdam: Elsevier.

Goldstein, I., A. Kopytov, L. Shen and H. Xiang (2022) *Synchronicity and Fragility*, Working Paper.

Grochulski, B. and Y. Zhang (2019) Optimal liquidity policy with shadow banking. *Economic Theory* 68(4), 967–1015.

Hanson, S. G., A. Shleifer, J. C. Stein and R. W. Vishny (2015) Banks as patient fixed-income investors. *Journal of Financial Economics* 117(3), 449–469.

Huang, J. (2018) Banking and shadow banking. *Journal of Economic Theory* 178, 124–152.

Kolotilin, A. and A. Wolitzky (2020). The economics of partisan gerrymandering. Working paper.

Li, Z. and K. Ma (2022) Contagious bank runs and committed liquidity support. *Management Science* 68(12), 8515–9218.

Liu, X. (2016) Interbank market freezes and creditor runs. *The Review of Financial Studies* 29(7), 1860–1910.

Liu, X. (2019) *Diversification and Systemic Bank Runs*, Working Paper.

Luck, S. and P. Schempp (2014) *Banks, Shadow Banking, and Fragility*, ECB Working Paper1726.

Manz, M. (2009) The optimal level of deposit insurance coverage. *Federal Reserve Bank of Boston,*Working Paper No 09.6.

Moreira, A. and A. Savov (2017) The macroeconomics of shadow banking. *The Journal of Finance* 72(6), 2381–2432.

Plantin, G. (2015) Shadow banking and bank capital regulation. *Review of Financial Studies* 28(1), 146–175.

Poszar, Z. (2011) Institutional cash pools and the triffin dilemma of the U.S. *Banking System. IMF Working Papers* No. 11/190.

Schilling, L. (2018) *Optimal Forbearance of Bank Resolution*, Becker Friedman InstituteWorking Paper 2018-15.

Voellmy, L. (2021) Preventing runs with fees and Gates. *Journal of Banking and Finance* 125, 106065.

Appendix

A.1. Proof of Lemma 4.1

First, define $f : [0, 1] \mapsto [0, 1]$, with $f(\vartheta) = 1 - p(D(\{j \in \mathcal{J} \mid \vartheta(j) \leqslant \vartheta\}))$. In words, $f(\vartheta)$ equals one minus the liquidation price in a situation where all banks whose share of retail deposits is below ϑ liquidate their portfolios. Since $f(\vartheta)$ is an increasing function mapping [0,1] into itself, we know from Tarski's fixed point theorem that the set of fixed points of $f(\vartheta)$ is nonempty and has a greatest element. Furthermore, denoting ϑ^{sr} as the greatest fixed point of $f(\vartheta)$, we have that $\vartheta \geqslant f(\vartheta)$ for any $\vartheta \geqslant \vartheta^{sr}$.

Next, I will show that the economy exhibits a run equilibrium that encompasses all banks with $\vartheta(j) < \vartheta^{sr}$. By condition (3), there exists a run equilibrium encompassing all banks with $\vartheta(j) < \vartheta^{sr}$ iff

$$\vartheta(j) < 1 - p(D(\{j \in \mathcal{J} \mid \vartheta(j) < \vartheta^{sr}\}) \quad \text{for all banks } j \in \mathcal{J} \text{ with } \quad \vartheta(j) < \vartheta^{sr}. \tag{9}$$

Rearranging the right-hand side of the inequality in condition (9) yields

$$1 - p(D(\{j \in \mathcal{J} \mid \vartheta(j) < \vartheta^{sr}\}) \leqslant 1 - p(D(\{j \in \mathcal{J} \mid \vartheta(j) \leqslant \vartheta^{sr}\}) = f(\vartheta^{sr}) = \vartheta^{sr},$$

which implies that condition (9) is fulfilled and there exists a run equilibrium encompassing all banks with $\vartheta(j) < \vartheta^{sr}$.

To complete the proof of Lemma 4.1, we need to show that there does not exist a run equilibrium encompassing banks with $\vartheta(j) \geqslant \vartheta^{sr}$. This can be shown with a proof by contradiction. Suppose there is a bank whose share of retail deposits equals $\tilde{\vartheta} \geqslant \vartheta^{sr}$ and suppose there exists a run equilibrium encompassing this bank. As is not hard to see, if the economy exhibits a run equilibrium encompassing a bank whose share of retail deposits is equal to $\tilde{\vartheta}$, then the economy must exhibit a run equilibrium encompassing all banks $j \in \mathcal{J}$ with $\vartheta(j) \leqslant \tilde{\vartheta}$. (This follows from the fact that, by condition (2), any bank that is susceptible to a run at some liquidation price p' will also be susceptible to a run at any liquidation price $p'' \leqslant p'$.) The economy exhibits a run equilibrium encompassing all banks with $\vartheta(j) \leqslant \tilde{\vartheta}$ iff

$$\vartheta(j) < 1 - p(D(\{j \in \mathcal{J} \mid \vartheta(j) \leqslant \tilde{\vartheta}\})) = f(\tilde{\vartheta}) \quad \text{for all banks } j \in \mathcal{J} \text{ with } \quad \vartheta(j) \leqslant \tilde{\vartheta}. \tag{10}$$

Since $\vartheta \geqslant f(\vartheta)$ for any $\vartheta \geqslant \vartheta^{sr}$, we have that $\tilde{\vartheta} \geqslant f(\tilde{\vartheta})$, so that condition (10) is violated, and we arrive at a contradiction.

We have thus shown that the economy does exhibit a run equilibrium encompassing all banks $j \in \mathcal{J}$ with $\vartheta(j) < \vartheta^{sr}$ and does not exhibit a run equilibrium encompassing any bank with $\vartheta(j) \geqslant \vartheta^{sr}$. It follows that a run encompassing all banks with $\vartheta(j) < \vartheta^{sr}$ is the largest run that is possible, which completes the proof of Lemma 4.1.

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

A.2. Proof of Proposition 4.1

Note first that, since the liquidation price cannot fall below p^m, we know from Condition (2) that a bank j whose share of retail deposits satisfies $\vartheta(j) \geqslant 1 - p^m$ will not be susceptible to runs. We can then proof Proposition 4.1 with a proof by contradiction. Suppose there is a structure of the banking system $\{D(j), \vartheta(j)\}_{j \in \mathcal{J}}$ that solves the planner's problem (4) and in which there is a bank $j \in \mathcal{J}$ with $0 < \vartheta(j) < \vartheta^{sr}$ issuing some strictly positive measure of deposits $D(j) > 0$. Suppose that, starting from this situation, we move all deposits of bank j into two other banks $j', j'' \in \mathcal{J}$. The reallocation must be such that the total amounts of retail and wholesale deposits remain unchanged, so that the planner's constraints remain satisfied. The reallocation must therefore satisfy conditions (11) and (12), which say that total retail and wholesale deposits, respectively, must remain constant:

$$\vartheta(j)D(j) = \vartheta(j')D(j') + \vartheta(j'')D(j'') \tag{11}$$

$$(1 - \vartheta(j))D(j) = (1 - \nu(j'))D(j') + (1 - \nu(j''))D(j'') \tag{12}$$

Consider now the reallocation of deposits characterized by:

$$(D(j'), \vartheta(j')) = \left(\frac{\vartheta(j)}{1 - p^m} D(j), 1 - p^m\right) \tag{13}$$

$$(D(j''), \vartheta(j'')) = \left(\left(1 - \frac{\vartheta(j)}{1 - p^m}\right) D(j), 0\right) \tag{14}$$

In words, bank j'' does not have any insured retail deposits, while bank j' has enough retail deposits to be shielded from runs. It is easy to verify that this reallocation of deposits satisfies conditions (11) and (12). Furthermore, since $\vartheta(j) < 1 - p^m$ (which is implied by $\vartheta(j) < \vartheta^{sr}$), both $D(j')$ and $D(j'')$ are strictly positive and strictly smaller than $D(j)$. Finally, as a result of the reallocation of deposits the fundamental value of assets liquidated in a systemic run decreases by $D(j) - D(j'') = D(j') > 0$, which implies that the initial allocation of deposits cannot be optimal. This completes the proof of Proposition 4.1.

A.3. Proof of Lemma 4.2

First, we can proof that $\vartheta_{CB} \geqslant \vartheta^{sr} \Rightarrow \vartheta_{CB} \geqslant 1 - p^m$ with a proof by contradiction. Suppose $\vartheta_{CB} \geqslant \vartheta^{sr}$ and $\vartheta_{CB} < 1 - p^m$. Since $1 - \vartheta(j) > p^m \equiv p(1)$ for all banks $j \in \mathcal{J}$, it follows from condition (3) that there exists a run equilibrium encompassing all banks $j \in \mathcal{J}$, which contradicts $\vartheta_{CB} \geqslant \vartheta^{sr}$. Next, we can proof $\vartheta_{CB} \geqslant 1 - p^m \Rightarrow \vartheta_{CB} \geqslant \vartheta^{sr}$ with a proof by contradiction as well. Suppose $\vartheta_{CB} \geqslant 1 - p^m$ and $\vartheta_{CB} < \vartheta^{sr}$. Since $\vartheta_{CB} < \vartheta^{sr}$, commercial banks are susceptible to systemic runs. However, since $1 - \vartheta_{CB} \geqslant p^m$, commercial banks are not susceptible to runs even if the liquidation price falls to the lowest possible level. Hence, commercial banks are not susceptible to systemic runs, which contradicts $\vartheta_{CB} < \vartheta^{sr}$ and completes the proof of Lemma A.3.

A.4. Proof of Proposition 4.2

Large parts of the proof of Proposition 4.2 are contained in the main text. First, it follows from the discussion in Subsection 4.2 that whenever the interval $[D_{SB}^{min}(\sigma), D_{SB}^{max}]$ is non-empty, systemic runs are prevented iff the size of the shadow banking sector is within said interval. It is easy to verify that the interval is non-empty iff $\sigma \geqslant (1 - p^m)(1 - \overline{\mu})$. Together with the fact that $D_{SB}^{min}(\sigma) = 0$ if $\sigma \geqslant (1 - p^m)$, this leads to items (i) and (ii) in Proposition 4.2.

Finally, if $\sigma < (1 - \overline{\mu})(1 - p^m)$, then we have $D_{SB}^{min}(\sigma) > D_{SB}^{max}$. By Proposition 4.1, minimizing the magnitude of systemic runs requires that commercial banks be not susceptible to runs, which requires $D_{SB} \geqslant D_{SB}^{min}(\sigma)$. Since $D_{SB}^{min}(\sigma) > D_{SB}^{max}$, shadow banks are susceptible to runs whenever commercial banks are not susceptible to runs. Minimizing the magnitude of runs thus requires to

https://doi.org/10.1017/S1365100522000694 Published online by Cambridge University Press

set the size of the shadow banking sector to the smallest size that satisfies $D_{SB} \geqslant D_{SB}^{min}(\sigma)$ which, evidently, is $D_{SB} = D_{SB}^{min}(\sigma)$. This leads to item (iii) in Proposition 4.2 and completes the proof of Proposition 4.2.

A.5. Proof of Proposition 5.2

Consider first the parameter region $\sigma < (1 - \overline{\mu})(1 - p^m)$. From Proposition 4.2, we know that preventing systemic runs entirely is not feasible in this region. Together with Proposition 5.1, this means that only type B equilibria can exist in this parameter region. To see that type B equilibria do indeed exist, consider a structure of the banking system where the share of retail deposits at all banks equals σ. Since the expected payoff at all banks is identical, depositors have no incentive to deviate. Furthermore, since $1 - \vartheta(j) = 1 - \sigma > p^m \equiv p(1)$ for all banks, we know from Condition (3) that there is a run equilibrium encompassing all banks, which completes the proof of item (i) in Proposition 5.2.

For the remainder of the proof, it will be useful to note that every structure of the banking system which is such that systemic runs do not occur constitutes a date 0 equilibrium. This follows from the fact that the expected payoff to both retail and wholesale deposits is identical at all banks if runs do not occur, so that depositors have no incentive to deviate. Consider the parameter region $\sigma \in [(1 - \overline{\mu})(1 - p^m), (1 - p^m))$. From Proposition 4.2, we know that preventing systemic runs is feasible in this region of σ. Since every structure of the banking system for which runs do not occur can be sustained as a date 0 equilibrium, it follows that type A equilibria do exist in this region of σ. To see why type B equilibria exist within this region as well, consider again a structure of the banking system where the share of retail deposits equals σ at all banks. Since $1 - \vartheta(j) = 1 - \sigma > p^m$ for all banks, there is a run equilibrium encompassing all banks. Since depositors have no incentive to deviate, this structure of the banking system constitutes an equilibrium as well, which completes the proof of item (ii) in Proposition 5.2.

Consider finally the region $\sigma \geqslant (1 - p^m)$. To see that type A equilibria exist in this region of σ, it suffices again to note that, by Proposition 4.2, it is feasible to prevent systemic runs in this region of σ, and every structure of the banking system for which systemic runs do not occur can be sustained as a date 0 equilibrium. To show that type B equilibria do not exist in this region of σ, we can proceed with a proof by contradiction. Suppose a type B equilibrium exists for $\sigma \geqslant (1 - p^m)$. By Proposition 5.1, all banks have an identical share σ of retail deposits in a type B equilibrium. Since $1 - \vartheta(j) = 1 - \sigma < p^m$ for all banks, no bank is susceptible to runs even if the liquidation price falls to the lowest possible level. This means that systemic runs do not occur, so that we are not in a type B equilibrium and have therefore arrived at a contradiction. This completes the proof of item (iii) in Proposition 5.2 and thus completes the proof of Proposition 5.2.

Cite this article: Voellmy L (2023). "The optimal distribution of insured and uninsured deposits across banks." *Macroeconomic Dynamics* **27**, 2252–2269. https://doi.org/10.1017/S1365100522000694