Journal of Financial Economics 161 (2024) 103929

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/finec

Systemic bank runs without aggregate risk: How a misallocation of liquidity may trigger a solvency crisis[☆]

Lukas Altermatt[a], Hugo van Buggenum[b], Lukas Voellmy[c,*]

[a] University of Essex, Wivenhoe Park, CO43SQ, Colchester, UK
[b] KOF-ETH Zürich, Leonhardstrasse 21, 8092, Zürich, Switzerland
[c] Swiss National Bank, Börsenstrasse 15, 8022, Zürich, Switzerland

ARTICLE INFO

Dataset link: Replication Package: Systemic Bank Runs without Aggregate Risk: How a Misallocation of Liquidity May Trigger a Solvency Crisis (Original data)

JEL classification:
E4
E5
G2

Keywords:
Fragility
Deposit freezes
Emergency liquidity

ABSTRACT

We develop a general equilibrium model of self-fulfilling bank runs. The key novelty is the way in which the banking system's assets and liabilities are connected. Banks issue loans to entrepreneurs who sell goods to households, which in turn pay for the goods by redeeming bank deposits. The return on bank assets is thus contingent on households being able to withdraw their deposits. In a run, not all households that wish to consume manage to withdraw, since part of banks' cash reserves end up in the hands of households without consumption needs. This misallocation of liquidity lowers revenues of entrepreneurs and bank asset returns, thereby rationalising the run. Interventions that restrict redemptions in a run can be self-defeating due to their negative effect on demand in goods markets. We show how runs can sometimes be prevented with combinations of deposit freezes and redemption penalties as well as with the provision of emergency liquidity.

1. Introduction

While systemic bank runs have long been an important topic in macro-financial economic research, we believe that the understanding of how systemic runs affect the real economy, how they may trigger solvency crises, and what this implies for the prevention of runs is still incomplete. Empirical observations show that systemic bank runs usually occur simultaneously with downturns in economic activity.[1] While a recession might increase the probability of a bank run, we argue that causality may also go the other way: a systemic run can lead to a misallocation of liquidity, which hinders economic activity and causes the downturn. The fact that returns on bank assets decrease in the downturn can then rationalise the run in the first place. Thus, financial crises and recessions may arise endogenously and reinforce each other.

In this paper, we provide new insights into the link between bank runs, misallocations of liquidity and bank insolvencies by incorporating bank deposits' role in the payment system into a general equilibrium banking model. The key feature of our model is that households redeem bank deposits to buy goods from entrepreneurs which in turn obtain loans from banks. The return on bank assets is thus endogenous and depends on aggregate demand, which itself depends on whether households with consumption needs (impatient households) can withdraw their deposits to buy goods. In a run, some households without consumption needs (patient households) redeem their deposits and store the money. Since part of banks' cash reserves end up in the hands of patient households, some impatient households cannot withdraw their deposits and, as a result, have to curtail their consumption. This misallocation of liquidity reduces aggregate demand in the goods market and leads some entrepreneurs to default on their loans obtained from banks. These losses on bank loans rationalise the run: patient households who keep their money invested receive less than the promised payout and are thus better off withdrawing early.

☆ Philipp Schnabl was the editor for this article. We thank him and the anonymous referee for comments and guidance that greatly improved the paper. We also thank Toni Ahnert, Mohammed Aït Lahcen, Aleksander Berentsen, Matt Darst (discussant), Hans Gersbach, Piero Gottardi, Allen Head, Hyeyoon Jung (discussant), Todd Keister, Cyril Monnet, Bruno Sultanum, Alex Vardoulakis and Randall Wright. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Swiss National Bank.

* Corresponding author.
E-mail addresses: l.altermatt@essex.ac.uk (L. Altermatt), hvanbuggenum@ethz.ch (H. van Buggenum), lukas.voellmy@snb.ch (L. Voellmy).

[1] See for example Reinhart and Rogoff (2008) or Gorton (1988).

https://doi.org/10.1016/j.jfineco.2024.103929
Received 5 May 2023; Received in revised form 16 August 2024; Accepted 18 August 2024
Available online 31 August 2024
0304-405X/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

A key implication of our model is that restrictions on deposit withdrawals may not eliminate run equilibria even if one assumes – as we do in our model – that there is no (fundamental) aggregate uncertainty. Therefore, linking the return on bank assets to the conditions in the real economy overturns the result from the literature following Diamond and Dybvig (1983) that suspension of deposit convertibility eliminates belief-driven runs if there are no fundamental shocks. In our model, deposit freezes do not eliminate runs because they do not resolve the misallocation of liquidity caused by runs. If banks impose a freeze, this prevents some impatient households from withdrawing, which reduces aggregate demand and thereby leads to lower revenues for bank borrowers and higher losses on bank loans. The anticipation of these losses in turn induces patient households to withdraw early. Here, it is important to point out the difference between *stopping* ongoing runs and *preventing* runs from starting in the first place. By definition, a full deposit freeze always stops a run since no further withdrawals are allowed. However, the negative consequences of a freeze for aggregate demand and banks' solvency mean that the incentive to run on the banks *before* the freeze is imposed may not be eliminated.

While the notion that restricting deposit withdrawals can impair economic activity is not new, explicitly modelling bank deposits' role in the payment system within a general equilibrium framework allows us to determine more precisely under which conditions the banking system is prone to runs due to this mechanism and how such runs may be prevented. We first show that the misallocation of liquidity created (ex post) by a run makes the banking system (ex ante) prone to runs unless the supply of assets unaffected by the shortfall in aggregate demand (e.g. government bonds) is high enough. We then analyse whether banks can prevent runs by imposing redemption penalties and partial deposit freezes once they observe a run. Such withdrawal restrictions can prevent runs if they satisfy three criteria: (i) it must be optimal for patient depositors to stop running once the restrictions are imposed; (ii) to limit the fall in aggregate demand, it must be optimal for impatient depositors to continue withdrawing when the restrictions are in place; (iii) the anticipation of withdrawal restrictions in case of a run must make it suboptimal for patient depositors to run ex ante. We show that for certain model parameters, no such policy exists, even if banks observe a run immediately.

Finally, we show that runs may be prevented by a central bank that buys illiquid assets from banks in case of a run. If such provision of emergency liquidity stabilises real spending of impatient households, then defaults by entrepreneurs are prevented, and the incentive to run on banks is eliminated. A sufficient condition to eliminate runs is that assets are purchased at face value and the intervention is not inflationary. However, this is only feasible if the central bank is sufficiently efficient in collecting loan repayments, since otherwise buying loans at face value will be loss-making for the central bank and inflationary. If assets are bought at a significant discount, some demand shortfall and thus some default by entrepreneurs is inevitable, in which case the intervention fails to eliminate runs. If the central bank instead purchases assets at face value despite the resulting rise in inflation, runs may or may not be eliminated depending on how inflation affects entrepreneurs' incentives to repay loans. Our result that runs can only be prevented if the central bank purchases assets at high enough prices can be seen as contradicting Bagehot (1873), who advocated for central banks that lend to solvent but illiquid banks at a high rate of interest.[2] Although Bagehot's policy prescription seems natural, in our model, the banking system's liquidity problem turns into a solvency problem only if the central bank charges a sufficiently high penalty on emergency liquidity.

While our paper is primarily a theoretical contribution, we also discuss the empirical relevance of the underlying mechanism. In our model, a bank run disrupts economic activity by hindering impatient depositors from obtaining the cash needed to buy goods. Therefore, our model might be most relevant for cash intensive economies. In this context, we provide anecdotal evidence suggesting that our mechanism was at work in historical banking panics in the U.S. and may also have played a role in the systemic run on Argentinian banks in 2001/02. We also discuss how a version of our mechanism could play out in economies in which payments can be made by direct deposit transfers without going through cash, for instance, if doubts about bank solvency in a crisis cause sellers to refuse payments in non-convertible deposits.

Related literature. Our model combines elements from Diamond and Dybvig (1983) and Aruoba et al. (2011), with the latter being itself based on the New Monetarist framework (Kiyotaki and Wright, 1989; Lagos and Wright, 2005). Similarly to Aruoba et al. (2011), entrepreneurs in our model use a neoclassical production technology (with capital and labour as inputs) to produce goods which they sell against fiat money to households. As in Diamond and Dybvig (1983), households face uncertainty regarding their future consumption needs, and they can form banking coalitions that provide insurance against idiosyncratic liquidity needs by pooling assets and issuing demand deposits. Our model shares many similarities with Geromichalos and Herrenbrueck (2022), who also introduce Diamond–Dybvig type liquidity shocks into a setting based on Aruoba et al. (2011). Rather than focusing on the role of banks as we do, Geromichalos and Herrenbrueck (2022) study the role of secondary asset markets (from which we abstract) in providing liquidity insurance.

Our paper contributes to a literature studying how banks can (or cannot) eliminate panic equilibria, in particular with regard to the role of deposit freezes. A key result of Diamond and Dybvig (1983) is that in the absence of aggregate uncertainty, runs can be prevented at no cost if banks fully freeze deposits after a certain number of depositors have withdrawn.[3] Up to this point, two main objections to this result have been raised in the theoretical literature. First, Engineer (1989) showed that in a riskfree Diamond–Dybvig setting with more periods, deposit freezes can give rise to run equilibria in which depositors withdraw preemptively out of fear of not being able to access their deposits when needed. Second, Ennis and Keister (2009) showed that Diamond and Dybvig's result hinges crucially on the assumption that banks can commit to freeze deposits even if this severely hurts some of their depositors ex post.[4] Our paper adds a third, distinct reason why deposit freezes may not eliminate panic equilibria in settings without aggregate uncertainty, even if banks can commit to any payout policy. Specifically, we highlight a general equilibrium effect of freezing deposits: by restricting access to cash for households that wish to consume, deposit freezes have a negative effect on aggregate demand, which – through its effect on bank asset returns – can rationalise the run in the first place.

The idea that early withdrawals negatively affect the long-run return on bank assets is shared by Andolfatto and Nosal (2020) and Kashyap et al. (2024), although for different reasons than in our paper. In Andolfatto and Nosal (2020), operating a bank entails a fixed cost, which means the return earned by those who remain invested falls when a bank downsizes due to early withdrawals. In Kashyap et al. (2024), early withdrawals negatively affect bank profitability, which

[2] Bagehot's dictum refers to central bank lending while we study the provision of liquidity through asset purchases. However, we show that the results are equivalent for collateralised lending.

[3] It is well known that full deposit freezes are generally not desirable in the presence of aggregate uncertainty, i.e. if either the investment technology is risky or liquidity needs of depositors are stochastic. Wallace (1990) provides a detailed discussion of full and partial deposit freezes in a model with random liquidity needs and a riskfree investment technology. Matta and Perotti (2024) show that full deposit freezes are generally not optimal when depositors' liquidity needs are known but the investment technology is risky.

[4] Relatedly, even if not explicitly focusing on deposit freezes, Keister and Mitkov (2023) show that it may be privately optimal for banks to refrain from imposing measures that stop a run if they expect to receive a government bailout when hit by a run.

in turn reduces bankers' incentive to monitor loans. While these papers follow different modelling approaches and do not deal with exactly the same questions as we do, it is clear that they have different implications regarding the prevention of runs. In particular, unlike in our model, putting an upper bound on early withdrawals will tend to protect the return on bank assets by limiting the potential decrease in banks' size and monitoring incentives, respectively.

Our paper is also related to a broader literature studying self-fulfilling systemic bank panics in general equilibrium. The previous literature on the topic, starting with the seminal contribution by Gertler and Kiyotaki (2015), has mostly focused on the case where general equilibrium effects of bank runs occur through the effect of runs on asset prices.[5] In these models, banks that are hit by runs liquidate assets, which depresses asset prices and has repercussions on the entire economy and financial system. The present paper does not feature fire sales but instead focuses on a different type of general equilibrium effect of bank runs, namely the misallocation of cash caused by runs. Two closely related papers are Robatto (2019) and Carapella (2015), both of which study general equilibrium effects of runs that are similar to our paper. In both papers, runs reduce the amount of liquid assets (cash or deposits) in the hands of those who wish to consume, which has deflationary effects and thus reduces banks' net worth. Similarly to our paper, the feedback effect from impaired liquidity provision by banks to asset prices can lead to self-fulfilling run equilibria. As both of these papers study endowment economies, they cannot speak to the mutually reinforcing effects of financial crises and real downturns, and they also do not examine the role of withdrawal restrictions in eliminating run equilibria.

Further, our contribution is part of a literature that studies the role of banks in providing liquidity insurance – and the fragility of these arrangements – within the New Monetarist framework. In our model, banks issue interest-paying demand deposits, which allows to insure depositors against random liquidity needs in a way that is reminiscent of Berentsen et al. (2007) and of Williamson (2012).[6] As we do in our paper, both Sanches (2018) and Andolfatto et al. (2020) combine the Lagos-Wright and Diamond-Dybvig frameworks. In Sanches (2018), a bank run impairs the use of bank deposits as a means of payment in DM trades, which can lead to a persistent negative effect on DM activity. Andolfatto et al. (2020) study a model in which banks and markets coexist and examine how the presence of markets affects banks' ability to provide liquidity insurance, a topic that is not the focus of our paper. Gu et al. (2023) highlight the inherent fragility of various aspects of banking in a wide variety of models, including one where banks act as a provider of a means of payment in a Lagos-Wright setting.

Finally, our paper is related to a theoretical literature studying the public sector's role in providing emergency liquidity in a run. The previous literature on the topic (e.g. Farhi and Tirole, 2020; Gorton and Huang, 2006; Martin, 2006; Rochet and Vives, 2004) has largely focused on the trade-off between preventing runs on the one hand and avoiding the creation of moral hazard on the other. Compared to these papers, we highlight banks' role in providing a means of payment and how this affects the way in which emergency liquidity needs to be provided in order to eliminate self-fulfilling panics.

Outline. The rest of this paper is structured as follows. Section 2 presents the environment. Section 3 briefly describes the planner's solution and the steady-state equilibrium without banks. Section 4 introduces banks and describes the steady-state banking equilibrium. Section 5 discusses bank runs, Section 6 examines whether redemption restrictions can stop or prevent runs, and Section 7 studies how government-provided emergency liquidity may eliminate runs. Finally, Section 8 discusses the empirical relevance of our model, and Section 9 concludes.

2. Environment

Time is discrete, indexed by $t = 0, 1, 2, ..$, and continues forever. Each period t is divided into two subperiods, called CM and DM.[7] The CM opens at the beginning of each period, and once it closes, the DM opens and remains open until the period ends. Both markets are competitive. There are two types of agents in the economy: a measure one of households and a measure n of entrepreneurs. Households are infinitely-lived, whereas entrepreneurs live for one period, from the CM until the CM next period, at which point they are replaced by a new generation of entrepreneurs. In the CM, a good x can be produced by households at linear disutility l, where one unit of l yields one unit of x. Good x cannot be stored, but it can be converted into capital k by young entrepreneurs one to one in order to produce another nonstorable good q in the DM. This DM good is produced according to

$$q = f(k, h), \tag{1}$$

where k denotes the amount of capital owned by the entrepreneur, h denotes his labour effort, and $f(k,h)$ has constant returns to scale (CRS). Capital fully depreciates after production. We assume that $f_k(k,h) > 0$, $f_{kk}(k,h) < 0$, $f_h(k,h) > 0$, $f_{hh}(k,h) < 0$, $f_{kh} > 0$ and $f(0,h) = f(k,0) = 0$. By inverting $f(k,h)$, we can rewrite it as

$$h = c(q, k), \tag{2}$$

where $c(q,k)$ is the amount of labour required to produce q units of the consumption good given k. From our assumptions on $f(k,h)$ it follows that $c(q,k)$ is homogeneous of degree one, and we have $c_q(q,k) > 0$, $c_{qq}(q,k) > 0$, $c_k(q,k) < 0$, $c_{kk}(q,k) > 0$ and $c_{qk}(q,k) < 0$.

During the DM, a fraction θ of households get utility $u(q)$ from consuming the good q. We call these agents *impatient* households. The remaining fraction $1-\theta$ get no utility from consumption during the DM, and we call these agents *patient* households. Households privately learn their type for the current period at the beginning of the DM. The realisation of types is i.i.d. across periods and households. In the CM, all households get utility $U(x)$ from consuming the CM good. The expected lifetime payoff of households is

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ U(x_t) - l_t + \theta u(q_t) \right\}, \tag{3}$$

where $U(x)$ and $u(q)$ are both strictly increasing, concave functions satisfying the Inada conditions. Additionally, we impose $u(0) = 0$.

Entrepreneurs get linear disutility from working during the DM and linear utility x from consuming during the CM when they are old. They have the same discount factor as households, such that the expected lifetime payoff of entrepreneurs born in period t equals

$$\mathrm{E}_t \left\{ -h_t + \beta x_{t+1} \right\}. \tag{4}$$

As is standard in the literature, we assume that agents are anonymous during the DM, that there is no record-keeping technology, and that households cannot credibly commit to any future payments. These assumptions rule out unsecured credit, which means households need liquid assets in order to purchase consumption goods from entrepreneurs during the DM. We assume that the only liquid asset in the sense that entrepreneurs can recognise it and distinguish it from counterfeited assets is intrinsically worthless fiat money m, which is

[5] See also Liu (2019) and Goldstein et al. (2020) for models of systemic bank panics with fire sales.

[6] In Berentsen et al. (2007), banks intermediate cash between agents with and without consumption needs. In our model, as in Williamson (2012), banks pool agents' cash and assets and issue demand deposits.

[7] This terminology follows standard conventions in the literature, whereby each time period is usually divided into two submarkets, a frictionless CM (which may stand for 'centralised market') and a DM (which may stand for 'decentralised market') in which trades need to be settled with fiat money.

issued by the government. Thus, households who want to consume in the DM are subject to the constraint

$$p_t q_t \le m_t, \tag{5}$$

where p_t denotes the price of the DM good in period t. We denote ϕ_t as the value of money during the CM of period t in terms of CM good, i.e. m_t units of fiat money buy $\phi_t m_t$ units of good x. We denote the gross inflation rate by $1+\pi_{t+1} \equiv \phi_t/\phi_{t+1}$.

Since young entrepreneurs do not have resources of their own, they need to borrow from households to invest in capital. We denote by ℓ_t nominal loans extended by households to entrepreneurs. To purchase k_t units of capital in the CM of period t, a young entrepreneur needs to take out a nominal loan of $\ell_t = k_t/\phi_t$. We denote $\tilde{i}_{\ell,t+1}$ as the net nominal interest rate on loans extended in period t, i.e. an entrepreneur receiving a nominal loan of ℓ_t in the CM of period t is due to repay $(1+\tilde{i}_{\ell,t+1})\ell_t$ units of money in the CM of period $t+1$. We assume entrepreneurs are able to commit to repayment, in the sense that they always repay their loan if they have the funds to do so. If entrepreneurs have insufficient funds to repay the loan in full, they will repay all they have. Notice, however, that entrepreneurs cannot be forced to work in the DM; in particular, they can always choose not to work at all in the DM, in which case they receive zero revenue and default on their entire loan.

In addition to fiat money M, the government issues nominal bonds B in the CM. A government bond issued in period t pays $1+i_{b,t+1}$ units of money in the CM of period $t+1$, such that the government's budget constraint writes

$$\phi_t(B_t + M_t) + \Delta_t = \phi_t(M_{t-1} + (1+i_{b,t})B_{t-1}), \tag{6}$$

where Δ_t denote lump-sum taxes (or subsidies if $\Delta_t < 0$) imposed on households in the CM of period t. We assume that the money supply M_t grows at a constant net rate μ, that the government targets a constant real debt level $\mathcal{B} = \phi_t B_t \ge 0$, and that lump-sum taxes Δ_t adjust such that the budget constraint holds given these targets.

For future reference, we denote $1+\iota_{t+1} \equiv (1+\pi_{t+1})/\beta$ as the Fisher rate, i.e. the nominal interest rate that fully compensates households for inflation and discounting.

3. Planner's problem and equilibrium without banks

Consider the problem of a social planner that maximises the expected lifetime utility of households and entrepreneurs, giving equal weight to all agents. We denote x_t and x_t^e as CM consumption of households and entrepreneurs, respectively, in period t. DM production per entrepreneur is denoted by q_t^e, while q_t denotes DM consumption per (impatient) household. The planner's problem then writes:

$$\max_{\{l_t,h_t,q_t,x_t,x_t^e\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \{U(x_t) - l_t + \theta u(q_t) + n(-h_t + x_t^e)\}$$
$$s.t. \quad l_t = x_t + n x_t^e + n k_t, \quad \theta q_t = n q_t^e, \quad h_t = c(q_t^e, k_t).$$

Inserting the constraints into the objective function and denoting $\kappa_t \equiv k_t/q_t^e$ as capital per DM good produced, we obtain the first-order conditions

$$U'(x^*) = 1, \tag{7}$$

$$u'(q^*) = c_q\left(1, \kappa^*\right), \tag{8}$$

$$1 = -c_k\left(1, \kappa^*\right), \tag{9}$$

where the (unique) first-best quantities are denoted with an asterisk (*).[8] Condition (7) equates households' marginal utility of CM consumption with the marginal disutility of CM production, (8) is the analogous condition for DM output, and (9) equates the marginal benefit of capital (through reduced hours worked in the DM) with the marginal disutility of producing capital.

Next, we will briefly describe the decentralised equilibrium in the economy without banks. In this and the next section, we will restrict attention to monetary steady-state equilibria, i.e. equilibria in which $\phi_t > 0$ for all t and in which all real quantities are constant.[9] In a monetary steady state, inflation equals the money growth rate, i.e. $\pi = \mu$, such that the Fisher rate equals $1+\iota = (1+\mu)/\beta$.

In the economy without banks, each household chooses its portfolio of money, bonds and loans (m_t, ℓ_t, b_t) in the CM. It is a standard result that households never choose to carry more money than what is needed to finance DM consumption if $\iota_{t+1} > 0$ (which is the relevant case for us), and we thus take it as given that $q_t = m_t/p_t$. Furthermore, we assume that each household holds a diversified loan portfolio, such that all households earn the same net return $i_{\ell,t+1}$ on their loans. We can then express households' problem as

$$\begin{aligned}\max_{\{x_t,m_t,\ell_t,b_t\}} U(x_t) - x_t + \theta u(m_t/p_t) &+ [\beta(1-\theta)\phi_{t+1} - \phi_t]m_t \\ &+ [\beta\phi_{t+1}(1+i_{\ell,t+1}) - \phi_t]\ell_t \\ &+ [\beta\phi_{t+1}(1+i_{b,t+1}) - \phi_t]b_t,\end{aligned} \tag{10}$$

where the last three terms capture the net benefit of carrying money, loans and bonds into the next CM (households only carry money into the next CM if they turn out to be patient).[10] From the first-order condition for m_t, we obtain

$$u'(q_t) = \rho_t\left(1 + \frac{\iota_{t+1}}{\theta}\right), \tag{11}$$

where $\rho_t \equiv \beta\phi_{t+1}p_t$ denotes the real price of the DM good in period t, and the term ι_{t+1}/θ represents the opportunity cost of holding money. Due to quasilinear preferences, interior solutions for ℓ_t and b_t are only possible if $i_{b,t+1} = i_{\ell,t+1} = \iota_{t+1}$: households' demand for an illiquid asset paying less (more) than ι_{t+1} would be zero (infinite). Market clearing thus implies that the equilibrium rate on loans and bonds equals the Fisher rate whenever these assets are in positive supply.

Next, entrepreneurs born in the CM of period t choose how much to invest in capital (k_t) and how much to work in the DM (h_t). Since investing k_t requires taking out a nominal loan k_t/ϕ_t, and since $h_t = c(q_t^e, k_t)$, we can express the entrepreneurs' problem as

$$\max_{\{k_t,q_t^e\}} \{-c(q_t^e,k_t) + \beta x_{t+1}^e\} \quad \text{s.t.} \quad x_{t+1}^e = \phi_{t+1}\left[p_t q_t^e - (1+\tilde{i}_{\ell,t+1})(k_t/\phi_t)\right]. \tag{12}$$

Using our definitions of κ_t and ρ_t, we obtain from the first-order conditions that

$$-c_k(1,\kappa_t) = \frac{1+\tilde{i}_{\ell,t+1}}{1+\iota_{t+1}}, \tag{13}$$

$$\rho_t = c_q(1,\kappa_t), \tag{14}$$

i.e. the optimal κ_t is decreasing in the loan rate $\tilde{i}_{\ell,t+1}$, and given κ_t, optimal DM production increases in ρ_t. We verify in Appendix A.2 that, given entrepreneurs' optimal investment choice, their equilibrium revenue is just sufficient to cover the cost of the loan and the disutility of working in the DM.

Since there is no uncertainty regarding the demand for DM good in the steady state, there are no defaults, which means that the effective loan rate equals the contractual rate, i.e. $i_{\ell,t+1} = \tilde{i}_{\ell,t+1}$. Market clearing for loans then requires $nk_t = \phi_t \ell_t$, while DM market clearing requires $nq_t^e = \theta q_t$. Finally, using (11)–(14) together with the fact that $\tilde{i}_{\ell,t+1} =$

[8] The optimal value of x^e is indeterminate since both the marginal utility of CM consumption of entrepreneurs as well as the marginal disutility of producing the CM good are equal to 1.

[9] There is always an equilibrium with $\phi_t = 0$ for all t in this class of models. In this equilibrium, $q = k = 0$ due to the lack of an accepted means of payment to settle DM transactions.

[10] Since the households' problem is standard for models based on Lagos and Wright (2005), we keep the exposition brief here and refer to Appendix A.1 for the details.

ι_{t+1} in equilibrium, we obtain that the steady-state levels of q and κ are determined by

$$\frac{u'(q)}{c_q(1,\kappa)} = 1 + \frac{\iota}{\theta}, \tag{15}$$

$$-c_k(1,\kappa) = 1. \tag{16}$$

Comparing (15)–(16) with (8)–(9), we see that the Friedman rule ($\iota = 0$) implements the first-best allocation, while DM output q is inefficiently low away from the Friedman rule. Furthermore, κ is always at its first-best level, i.e. a given q is produced with the socially optimal mix of capital and labour.

4. Banking equilibrium

To overcome the uncertainty about their consumption preferences, households can form coalitions during the CM, which we call banks, à la Diamond and Dybvig (1983). Banking coalitions maximise the expected utility of their participating households, which we call depositors, with each depositor depositing an identical amount in the bank when it is formed. Banks act as price takers in the sense that they take equilibrium loan and bond rates, the value of money, and the price of the DM good as given. Banks exist for one period, from the CM until the CM next period, at which point they are dissolved and replaced by a new set of banks.

In this section, we solve for steady-state banking equilibria. In the subsequent sections, we then investigate whether the steady-state equilibria we find are prone to unexpected runs.

Banks issue deposits in the CM when they are formed and invest the proceeds in cash (m_t^b), loans (ℓ_t^b) and government bonds (b_t^b), where we use the superscript b to denote portfolio choices made by banks. We denote by $a_t^b \equiv \ell_t^b + b_t^b$ a bank's total holdings of illiquid assets. Since expected loan defaults are zero in the steady state, the nominal return on loans and bonds must be the same in equilibrium, i.e. $i_{\ell,t+1} = i_{b,t+1} \equiv i_{t+1}$ for all t. Thus, when we refer to the interest rate, this covers both of these rates.[11] A bank formed in the CM of period t promises to pay d_t^I units of money to impatient depositors in the following DM and d_{t+1}^P units of money to patient depositors in CM next period. We have $d_t^I = p_t q_t$, i.e. choosing the DM payout d_t^I is equivalent to choosing the DM consumption q_t of impatient households.[12] The bank's problem can then be expressed as

$$\max_{m_t^b, a_t^b, d_t^I, d_{t+1}^P} \theta u(d_t^I/p_t) + \beta(1-\theta)\phi_{t+1} d_{t+1}^P - \phi(m_t^b + a_t^b)$$

subject to:

$$m_t^b \geq \theta d_t^I, \qquad (\beta\phi_{t+1}\zeta_t) \qquad \text{liquidity const.} \tag{17}$$

$$m_t^b + a_t^b(1+i_{t+1}) \geq \theta d_t^I + (1-\theta) d_{t+1}^P, \qquad (\beta\phi_{t+1}\xi_t) \qquad \text{solvency const.} \tag{18}$$

$$d_{t+1}^P \geq d_t^I, \qquad (\beta\phi_{t+1}(1-\theta)\psi_t) \qquad \text{patient IC const.} \tag{19}$$

$$u(q_t) - u(0) \geq \beta\phi_{t+1} d_{t+1}^P, \qquad \text{impatient IC const.} \tag{20}$$

where the terms in brackets denote the Lagrange multipliers on the constraints. Condition (17) states that the bank must hold enough money to make payments to early withdrawers. Condition (18) states that the total value of the bank's assets must be at least as high as the total promises it makes to depositors. Condition (19) is an incentive compatibility (IC) constraint for patient depositors, which states that DM payouts cannot exceed CM payouts. Finally, condition (20) is an IC constraint for impatient depositors, which states that the difference in utility from consuming q_t in the DM instead of nothing should exceed the utility impatient depositors obtain from withdrawing in the CM. We will ignore this constraint for now and check later whether it binds.

Before moving on, we want to briefly elaborate on the nature of liquidity insurance that banks provide in our model. Different to the Diamond and Dybvig (1983) model, the purpose of liquidity insurance is not to smooth consumption between impatient and patient households. Instead, liquidity insurance ensures that households receive money in the DM if and only if they need it. While in autarky all households hold some money to self-insure against the risk of being impatient, banks efficiently distribute the available money to those households who need it, allowing impatient depositors to consume more than they would in autarky. In our setup, 'full insurance' against liquidity risk means that the real amount of money that impatient households receive from banks in the DM equals the amount that households would bring into the DM if they lived in autarky and knew their type beforehand.[13] Note also that the sole reason why banks hold illiquid assets in our model is to satisfy patient depositors' IC constraint. If types were observable, banks could just invest in money and distribute it to impatient depositors in the DM; but since types are private information, banks must hold illiquid assets to make payments to patient depositors that are large enough to satisfy incentive compatibility.[14]

We refer to Appendix A.3 for the derivation of the optimality conditions of the bank's problem. For the Lagrange multipliers, we obtain

$$\zeta_t = \frac{i_{t+1}(1+\iota_{t+1})}{1+i_{t+1}}, \quad \xi_t = \frac{1+\iota_{t+1}}{1+i_{t+1}} \quad \text{and} \quad \psi_t = \frac{\iota_{t+1} - i_{t+1}}{1+i_{t+1}}, \tag{21}$$

i.e. the liquidity constraint binds unless the nominal rate is zero, the solvency constraint always binds (which just means that banks make zero profits) and patient depositors' IC constraint binds unless the nominal rate equals the Fisher rate. Intuitively, the latter is the case because if assets pay the Fisher rate, buying additional assets in order to increase CM payouts and relax patient depositors' IC constraint entails no opportunity cost for banks. Further, we obtain that DM consumption satisfies

$$u'(q_t) = \rho_t \left[1 + \frac{\theta\iota_{t+1}(1+i_{t+1}) + (1-\theta)(\iota_{t+1} - i_{t+1})}{\theta(1+i_{t+1})}\right], \tag{22}$$

and banks' asset demand schedule is

$$m_t^b \geq \theta p_t q_t, \qquad \text{with equality if } i_{t+1} > 0, \tag{23}$$

$$m_t^b + a_t^b(1+i_{t+1}) \geq p_t q_t, \qquad \text{with equality if } i_{t+1} < \iota_{t+1}. \tag{24}$$

Condition (23) implies that banks only hold excess cash if the nominal interest rate is zero, and condition (24) means that DM payouts cannot exceed the bank's nominal portfolio return, which results from patient depositors' IC constraint.

The presence of banks does not change entrepreneurs' optimisation problem relative to the economy without banks, such that the equilibrium conditions (13)–(14) are still valid.

[11] From Section 3, we know that households are only willing to hold illiquid assets outside of banks if they pay the Fisher rate. As we will see, banks might be willing to hold illiquid assets at lower rates, but if $i_{\ell,t+1} \neq i_{b,t+1}$, they would strictly prefer to hold whichever asset pays the higher interest rate. Recall also that asset returns cannot exceed the Fisher rate since demand from households for such an asset would be infinite.

[12] We assume here that households hold no money outside of banks; this is without loss of generality since households would never find it optimal to do so.

[13] In other words, full liquidity insurance is equivalent to eliminating uncertainty about types. This is different from the Diamond–Dybvig model, where liquidity insurance entails a redistribution between types. See Voellmy (2024) for a detailed discussion of the nature of liquidity insurance in different banking models.

[14] Our finding that incentive compatibility for patient households may limit banks' ability to provide liquidity insurance is similar to Peck and Setayesh (2023). In their model, banks can only provide optimal liquidity insurance if households deposit a large enough share of their endowment in the bank.

Market clearing in the markets for loans, bonds and DM good, respectively, requires

$$nk_t = \phi_t(\ell_t + \ell_t^b), \quad B_t = b_t + b_t^b, \quad \text{and} \quad nq_t^e = \theta q_t.$$

Combining (13), (14) and (22), and using the fact that there are no defaults in the steady state (which means $\tilde{i}_{\ell,t+1} = i_{t+1}$), we obtain that the steady-state values of q and κ satisfy

$$\frac{u'(q)}{c_q(1,\kappa)} = 1 + \frac{\theta\iota(1+i) + (1-\theta)(\iota - i)}{\theta(1+i)}, \tag{25}$$

$$-c_k(1,\kappa) = \frac{1+i}{1+\iota}. \tag{26}$$

For given nominal rates i and ι, this system has a unique solution. Eq. (26) shows that κ decreases in i. Inspecting Eq. (25) reveals that changes in i have two opposing effects on DM output q. On the one hand, an increase in i allows banks to pay a higher deposit rate, thereby decreasing the effective cost of liquidity, which taken by itself has a positive effect on q. On the other hand, the fall in κ caused by an increase in i leads to an increase in the real DM price $\rho = c_q(1,\kappa)$ (see (14)), which taken by itself has a negative effect on q.

Next, we denote by $\mathcal{A}$ the real aggregate asset supply, which equals the aggregate capital stock plus real government debt:

$$\mathcal{A} \equiv \phi_t a_t = nk + \mathcal{B}. \tag{27}$$

Combining (27) with the representative bank's asset demand schedule (23)–(24), and using the fact that all assets are held by banks if $i < \iota$, we obtain that asset market clearing requires:

$$\mathcal{A} \begin{cases} \geq (1-\theta)\rho q & \text{if} \quad i = \iota \\ = \dfrac{1+\iota}{1+i}(1-\theta)\rho q & \text{if} \quad i \in (0,\iota) \\ \leq (1+\iota)(1-\theta)\rho q & \text{if} \quad i = 0. \end{cases} \tag{28}$$

Definition 1. A stationary monetary equilibrium (SME) in the banking economy is given by $(q, \kappa, \rho, i, \mathcal{A})$ satisfying (14) and (25)–(28).

The existence proof for SME is given in Appendix A.4:

Proposition 1. *A stationary monetary equilibrium in the banking economy exists.*

In the following, we will distinguish between three equilibrium cases: (i) *full liquidity insurance equilibria (FLI)*, defined as equilibria in which $i = \iota$; (ii) *zero lower bound equilibria (ZLB)*, defined as equilibria in which $i = 0$ and (iii) *partial liquidity insurance equilibria (PLI)*, defined as equilibria in which $i \in (0,\iota)$. Roughly speaking, condition (28) states that an FLI equilibrium exists if the aggregate asset supply is plentiful, a ZLB equilibrium exists if assets are scarce, and a PLI equilibrium exists if the asset supply is within an intermediate range. In a separate paper (Altermatt et al., 2024), we discuss in more detail the steady-state results in a model similar to this one, and we show that all three equilibrium cases may coexist for certain parameters. However, this coexistence is immaterial for the purposes of the present paper, and we will only very briefly characterise the equilibrium cases here before moving on to bank runs.

In an FLI equilibrium, banks are saturated with assets, such that asset prices fall to the point where assets pay the Fisher rate. We have $\kappa = \kappa^*$ in an FLI equilibrium, and DM output q satisfies $u'(q)/\rho = 1 + \iota$, i.e. q is higher than in an economy without banks and is unaffected by the probability of being impatient (θ). In an FLI equilibrium, banks fully insure depositors against liquidity risk in the sense that DM consumption is the same as it would be in an economy in which households know their type beforehand. Consider next ZLB equilibria, in which banks' demand for assets drives the interest rate down to the zero lower bound. Then, banks are indifferent between holding illiquid assets and cash, and the nominal deposit return is zero. DM output in a ZLB equilibrium satisfies $u'(q)/\rho = 1 + (\iota/\theta)$, which is the same condition as in the economy without banks. However, the depressed loan rate implies $\kappa > \kappa^*$ in a ZLB equilibrium, which in turn implies that ρ is lower and q is higher than in the economy without banks. Finally, PLI equilibria represent an intermediate case between FLI and ZLB equilibria.

Let us now return to the question whether the IC constraint for impatient depositors (condition (20)) is indeed fulfilled in equilibrium, as we have assumed so far. We know from (21) that in ZLB and PLI equilibria, the IC constraint for *patient* depositors binds, which makes it straightforward to show that the IC constraint for *impatient* depositors is slack. In an FLI equilibrium, the IC constraint for impatient depositors is satisfied as long as

$$\phi a^b \leq \bar{\mathcal{A}} \equiv (1-\theta)u(q), \tag{29}$$

i.e. as long as banks' holdings of illiquid assets do not exceed the threshold $\bar{\mathcal{A}}$. For FLI equilibria to exist for some parameters, the lower bound on $\mathcal{A}$ for which an FLI equilibrium can exist (see (28)) needs to be below the upper bound in (29). This is the case if

$$\frac{1}{1+\iota}u'(q)q \leq u(q),$$

which is always strictly satisfied given the properties of $u(q)$. From here on, we will assume that banks hold all illiquid assets if $\mathcal{A} \leq \bar{\mathcal{A}}$ and that they hold $\bar{\mathcal{A}}$ if $\mathcal{A} > \bar{\mathcal{A}}$.

5. Bank runs

We now study whether steady-state banking equilibria are prone to self-fulfilling runs. We treat runs as zero-probability events, which therefore have no effect on the steady-state allocations studied previously. From now on, we will use subscript S to denote the steady-state values of variables.

A run denotes a situation where all depositors, both patient and impatient, rush to the banks in order to withdraw their deposits in the DM. Following Diamond and Dybvig (1983) and much of the banking literature, we assume sequential service, i.e. depositors arrive at their bank in random order in the DM and need to be paid out on the spot. Depositors who wish to redeem in the DM form a queue, with each depositor being assigned each place in the queue with identical probability. Patient depositors run if and only if they expect that withdrawing in the DM gives them a strictly higher payout compared to not withdrawing.[15] We will say that the banking system is *fragile* if steady-state DM payouts (d_S^I) are strictly higher than CM payouts in a situation where all depositors run.

In a run, some cash ends up in the hands of patient depositors, who will not spend it until the next CM. This means that impatient depositors hold less cash in aggregate than in the steady state. We denote by M^{RE} the total cash paid out to impatient depositors in the DM in a run, with $0 < M^{RE} \leq \theta d_S^I$, where θd_S^I equals the total cash paid out to impatient depositors in the steady state.

If a run occurs, each entrepreneur will either repay his loan in full or not produce at all in the DM and default on his entire loan. To see this, recall that if an entrepreneur produces in the DM but fails to repay his loan in full, he will hand over his entire revenue to the creditors; since production entails a utility cost, this leads to a strictly negative payoff for the entrepreneur, meaning the entrepreneur would have been better off not producing at all.

We denote by χ the share of entrepreneurs that choose to produce in the DM and thus repay their loan in a run. Active (i.e. producing) entrepreneurs choose their supply, q^e, to maximise real profits before loan repayments, $\rho q^e - c(q^e, k_S)$. Supply by an active entrepreneur is

[15] We assume patient depositors do not redeem in the DM (they 'stay home') in case of indifference.

therefore set according to $q^e = q^e(\rho) \equiv c_q^{-1}(\rho, k_S)$, which implies that $q^e(\rho)$ is strictly increasing in ρ.

Market clearing in the DM requires that aggregate real spending on the DM good by impatient depositors equals aggregate real revenues of entrepreneurs:

$$\frac{1+\pi_S}{1+\pi}\frac{M^{RE}}{\theta d_S^I} = \frac{\chi \rho q^e(\rho)}{\rho_S q^e(\rho_S)}. \tag{30}$$

The LHS in Eq. (30) equals impatient depositors' aggregate real cash holdings in a run relative to the steady state, and the RHS equals entrepreneurs' aggregate real revenues relative to the steady state.

As long as no additional money is injected in a run, the inflation rate is not affected by the run, and we obtain a particularly simple expression for the share of active entrepreneurs:

Proposition 2. *If the amount of money in circulation does not change when a run occurs, the real purchasing power of money remains unchanged ($\pi = \pi_S$). Then,*

(i) the real price of the DM good remains at the steady-state level ($\rho = \rho_S$), and
(ii) the share of active entrepreneurs equals $\chi = \frac{M^{RE}}{\theta d_S^I}$.

Proof. The fact that $\pi = \pi_S$ if the money supply does not deviate from its steady-state value follows from the fact that steady-state inflation is determined through the money growth rate, and the economy reverts to the steady state after a run. To see why ρ cannot deviate from ρ_S, note first that entrepreneurs make zero profits in the steady state, in the sense that their payoff-maximising DM labour effort given k_S and ρ_S yields them a payoff of zero (see Appendix A.2). If $\pi = \pi_S$, the real indebtedness of entrepreneurs stays the same, which implies that ρ cannot fall below ρ_S: a decrease in ρ shifts the payoff schedule downwards, such that $\rho < \rho_S$ implies a strictly negative payoff for entrepreneurs for any production level $q^e > 0$. Therefore, $\rho < \rho_S$ implies zero production in the DM, which violates market clearing (30). Next, $\rho > \rho_S$ cannot be consistent with market clearing (30) either, as it would lead to an increase in the aggregate supply of the DM good while the aggregate demand for the DM good decreases. Finally, given $\rho = \rho_S$ and $\pi = \pi_S$, item (ii) in Proposition 2 follows immediately from (30). ■

For banks, the effective gross nominal return on loans in a run equals $\chi(1+i_S)$. Notice that the return to an individual bank's portfolio depends on the redemption behaviour of depositors at all banks, making runs inherently systemic events. We want to highlight here that the link between the occurrence of a run and the return on illiquid assets captured by item (ii) in Proposition 2 is the main innovation in our model compared to the existing literature.[16]

Consider now the more general case where a run may affect the price level and hence the real price of the DM good, ρ. This is only relevant for the analysis in Section 7, where we study the injection of emergency liquidity by the government.

Proposition 3. *Given any real price ρ of the DM good in a run, the share of active entrepreneurs equals*

$$\chi = \min\left\{\frac{\alpha(\rho)}{\alpha(\rho_S)}\frac{M^{RE}}{\theta d_S^I}, 1\right\}, \tag{31}$$

where

$$\alpha(\rho) = \frac{\rho q^e(\rho) - c(q^e(\rho), k_S)}{\rho q^e(\rho)} \tag{32}$$

denotes the capital share, i.e. the share of an active entrepreneur's revenue left after compensating the entrepreneur for his labour disutility of production.

Proof. An entrepreneur is willing to produce and repay his loan if and only if

$$\max_{q^e}\{\rho q^e - c(q^e, k_S)\} \equiv \rho q^e(\rho) - c(q^e(\rho), k_S) = \rho\alpha(\rho)q^e(\rho) \geq \beta\frac{1+i_S}{1+\pi}k_S, \tag{33}$$

i.e. if and only if the maximised real profits before repayment weakly exceed the real debt burden. Exploiting that (33) holds with equality in the steady state (see Appendix A.2), we find

$$\frac{\rho\alpha(\rho)q^e(\rho)}{\rho_S\alpha(\rho_S)q^e(\rho_S)} \geq \frac{1+\pi_S}{1+\pi}, \quad \text{with equality if } \chi < 1, \tag{34}$$

i.e. to incentivise entrepreneurs to produce in the DM, their real revenue left after compensating them for their labour disutility cannot decrease by more than the reduction in the real debt burden caused by an increase in inflation. Combining (30) and (34) then leads to the result in Proposition 3. ■

Whether the banking system is fragile ex ante depends on the banks' and the government's ex post reaction to runs. In Section 6, we study withdrawal restrictions imposed by banks, which can be interpreted as partial deposit freezes and redemption penalties. In Section 7, we widen the scope of our analysis by also considering public provision of emergency liquidity, where the government stands ready to purchase illiquid assets from banks. In line with the bulk of the bank run literature, we assume the banking system reacts to runs as a single, consolidated entity.[17] The speed at which the queue is served is the same at all banks, i.e. whenever a bank has served some fraction of queuing depositors in the DM, all other banks will have served the same fraction of their queue.

We assume for the remainder of the paper that all banks and the government realise simultaneously that a run is underway after a fraction $\lambda \in [0, \theta]$ of depositors have withdrawn in the DM, at which point banks may jointly impose redemption restrictions or the government starts to inject emergency liquidity. Since the share of impatient depositors equals θ, banks always know that a run is going on if more than a fraction θ of depositors wish to redeem in the DM. $\lambda = 0$ means that the run is immediately spotted, i.e. after a measure zero of patient depositors have withdrawn. Note also that since banks cannot liquidate assets in the DM, they necessarily implement a full deposit freeze once they run out of cash unless the government provides emergency liquidity.

In case banks impose a partial freeze or redemption penalties after realising that a run is underway, depositors who were not among the first λ to show up can choose whether they still want to redeem in the DM or whether they want to leave the queue and be paid out in the CM instead. If patient depositors leave the queue once the redemption restrictions are in place, we say that these measures *stop* runs. If the banking system is not fragile given that banks impose redemption

[16] Note that we do not allow entrepreneurs to renegotiate their debt contracts. If entrepreneurs could renegotiate, Proposition 2 would not necessarily hold anymore, as entrepreneurs might be willing to sell DM goods at lower prices if their repayment burden is reduced. While this could mitigate the macroeconomic effects of a bank run, it would not help the banking system in stopping or preventing runs, since as long as entrepreneurs repay less than was initially promised, run incentives for patient depositors remain intact.

[17] One interpretation of this is that banks jointly commit to reacting in a certain way if they realise that a run is underway. Ennis and Keister (2009, 2010) interpret the concerted action by banks as the result of a centralised banking authority stepping in once a systemic run has started, where the banking authority could be regarded as a reduced-form representation of the banking system together with the relevant regulatory agencies. Importantly, unlike the banking authority in Ennis and Keister (2009, 2010), the consolidated banking system does not suffer from limited commitment in our model.

restrictions and/or the government provides emergency liquidity after observing a run (off the equilibrium path), we say that these measures *prevent* runs.

6. Deposit freezes and redemption penalties

We first consider deposit freezes and then introduce redemption penalties. Suppose banks can freeze any fraction $1-\eta \in (0,1]$ of deposits upon realising that a run is underway. If deposits are (partially) frozen, depositors can only redeem a fraction η of their deposit in the DM, thus receiving a DM payout of ηd_S^I, while the remaining part of the deposit is locked in until the CM. In the CM, depositors are paid out pro-rata, i.e. if depositors who did not withdraw anything in the DM receive some amount d in the CM, then depositors who redeemed fraction η of their deposit in the DM receive $(1-\eta)d$ in the CM. The standard full deposit freeze studied by Diamond and Dybvig (1983) and others corresponds to $\eta = 0$ and $\lambda = \theta$.

Recall that at the point in time when banks impose a deposit freeze, a fraction λ of depositors have already withdrawn. Denoting ω as the share of depositors that redeem in the DM *after* a partial freeze with $\eta > 0$ has been imposed, the banks' liquidity constraint implies

$$\omega \leq \overline{\omega}(\eta) \equiv \min\left\{1-\lambda, \frac{m_S^b - \lambda d_S^I}{\eta d_S^I}\right\}. \tag{35}$$

For future reference, we define

$$\overline{\eta} \equiv \min\left\{\frac{m_S^b - \lambda d_S^I}{(1-\lambda)d_S^I}, 1\right\} \quad \text{and} \quad \overline{\overline{\eta}} \equiv \min\left\{\frac{\overline{\eta}}{\theta}, 1\right\}, \tag{36}$$

such that $\overline{\omega}(\overline{\eta}) = 1-\lambda$ and $\overline{\omega}\left(\overline{\overline{\eta}}\right) = \theta(1-\lambda)$. That is, after realising that a run is underway, banks' cash reserves are just sufficient to convert a fraction $\overline{\eta}$ of all remaining deposits into cash in the DM; and cash reserves are just sufficient to convert a fraction $\overline{\overline{\eta}} > \overline{\eta}$ of all remaining deposits held by *impatient* depositors into cash.

Suppose now that, in addition to partially freezing deposits, banks can impose a penalty on DM redemptions, which we denote by σ. Depositors who redeem the allowed fraction η of their deposit in the DM then forgo a fraction $\sigma \in [0, 1-\eta]$ of their deposit. That is, if depositors who do not redeem in the DM receive some amount d in the CM, those redeeming fraction η of their deposit receive $(1-\eta-\sigma)d$ in the CM.

Banks' solvency constraint in case of a run writes

$$m_S^b + (1+i_S)(b_S^b + \chi \ell_S^b) = \lambda d_S^I + \omega(\eta d_S^I + (1-\eta-\sigma)d_R^P) + (1-\lambda-\omega)d_R^P, \tag{37}$$

where d_R^P denotes the CM payout to depositors who do not withdraw in the DM. From (37), we obtain that CM payouts after a run equal

$$d_R^P(\chi, \omega; \eta, \sigma) = \frac{[m_S^b - (\lambda + \eta\omega)d_S^I] + (1+i_S)(b_S^b + \chi \ell_S^b)}{1-\lambda-\omega(\eta+\sigma)}. \tag{38}$$

The numerator in (38) equals the total resources of a bank in the CM given that a run occurred in the DM; the first term in the numerator equals left-over cash not paid out in the DM, and the second term equals the proceeds from a bank's portfolio of government bonds and loans. The denominator is the measure of outstanding deposits in the CM.

6.1. Deposit freezes

In this subsection, we focus on pure deposit freezes without redemption penalties, i.e. we set $\sigma = 0$. Findings in the previous literature suggest that deposit freezes are effective in preventing runs in economies without aggregate uncertainty and with full commitment. We show that this is not the case once general equilibrium effects of deposit freezes are taken into account.

We start by noting that a partial deposit freeze with $\eta \in (0,1)$ cannot stop a run that has already started, in the following sense: if redeeming the entire deposit is patient depositors' best response, then so is redeeming any fraction $\eta > 0$ of the deposit. To see this, note that redeeming the allowed fraction η is patient depositors' best response if and only if $\eta d_S^I + (1-\eta)d_R^P(\cdot) > d_R^P(\cdot)$, which is equivalent to $d_S^I > d_R^P(\cdot)$.

To understand whether the banking system is fragile, we need to determine how many entrepreneurs default in a run. Consider deposit freezes with $\eta \geq \overline{\eta}$, in which case banks pay out their entire cash holdings to withdrawing depositors in a run. Then, a fraction λ of depositors manages to withdraw in full in a run, a fraction $\overline{\omega}(\eta)$ manages to withdraw a share η of their deposit, while the remaining fraction $1-\lambda-\overline{\omega}(\eta)$ of depositors cannot redeem anything in the DM. By a law of large numbers, a share θ of the depositors who manage to withdraw in the DM will be impatient, which implies that the total cash paid out to impatient depositors equals θm_S^b. By Proposition 2, we then have

$$\chi = \frac{m_S^b}{d_S^I} \in (0,1) \tag{39}$$

for $\eta \geq \overline{\eta}$. Combining (38) and (39), we obtain that the banking system is fragile under any deposit freeze satisfying $\eta \geq \overline{\eta}$ if and only if

$$d_S^I > d_R^P\left(\frac{m_S^b}{d_S^I}, \overline{\omega}(\eta); \eta, 0\right) \quad \Leftrightarrow \quad d_S^I > m_S^b + (1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right). \tag{40}$$

In words, the banking system is fragile if the steady-state DM payouts are higher than the value of banks' portfolios after taking into account loan defaults caused by a run. Notice that a freeze with $\eta < \overline{\eta}$ would exacerbate incentives to run compared to a freeze with $\eta \geq \overline{\eta}$ since the total cash paid out to impatient depositors would then be strictly below θm_S^b, such that χ would be lower than in (39). Therefore, condition (40) is a sufficient condition for the banking system to be fragile under any deposit freeze $\eta \in [0,1]$, which leads to the following result:

Proposition 4. *For any $\lambda \in [0,\theta]$ and any deposit freeze $\eta \in [0,1]$, the banking system is fragile in all zero lower bound (ZLB) and partial liquidity insurance (PLI) equilibria, as well as in full liquidity insurance (FLI) equilibria in which the real aggregate asset supply satisfies $\mathcal{A}_S < \min\{\bar{\mathcal{A}}, \bar{\mathcal{A}}^F\}$, where $\bar{\mathcal{A}}^F \equiv (1-\theta)q_S(\rho_S + \theta\kappa^*)$.*

Proof. Consider first ZLB and PLI equilibria. In these equilibria, we have $d_S^I = d_S^P = m_S^b + (1+i_S)(b_S^b + \ell_S^b)$. Since $m_S^b/d_S^I < 1$, it follows immediately from condition (40) that the banking system is fragile in ZLB and PLI equilibria. Consider next FLI equilibria. In FLI equilibria, we have $\theta d_S^I = m_S^b$. Substituting this and $d_S^I = p_S q_S$ into condition (40) and rearranging yields the condition $(1-\theta)p_S q_S > (1+\imath)(b_S^b + \theta \ell_S^b)$. Substituting $\rho \equiv [\phi/(1+\imath)]p$, $\phi \ell^b = \theta q \kappa$, $\mathcal{A} \equiv \phi(b^b + \ell^b)$ and using the fact that $\kappa = \kappa^*$ in an FLI equilibrium yields that, assuming banks hold all assets in the economy (which is the case if $\mathcal{A}_S \leq \bar{\mathcal{A}}$), they are fragile iff $\mathcal{A}_S \leq \bar{\mathcal{A}}^F$.[18] ■

The fragility of the banking system in ZLB and PLI equilibria is closely related to the fact that the IC constraint for patient depositors binds in these equilibria. Even slight losses on banks' investments then cause CM payouts to fall below DM payouts, making it optimal for patient depositors to run. In FLI equilibria, patient depositors receive strictly higher payouts than impatient depositors in the steady state, such that banks have a buffer to absorb a certain amount of losses on their loans.

A remarkable result of Proposition 4 is that if banks' buffer to absorb losses is small, deposit freezes can neither prevent runs nor stop them from starting, no matter how quickly banks react to runs and what fraction of deposits they freeze. Even if banks can impose a (partial) deposit freeze immediately after a run has started, the fact that impatient depositors can withdraw less cash due to the freeze

[18] If $\mathcal{A} > \bar{\mathcal{A}}$ then not all assets in the economy are held by banks. Whether banks are fragile then depends additionally on how illiquid assets are allocated between households and banks; the larger the share of government bonds (loans) in banks' portfolios is, the less (more) likely it is that banks are fragile.

makes it optimal for patient depositors to run. This latter point is the major difference between our paper and other papers in the literature in which the return on bank assets can be protected by freezing deposits, as this prevents banks from incurring liquidation losses. In our model, the return on bank assets is tied to the conditions in the real economy; a deposit freeze implies that less cash ends up in the hands of impatient depositors, which hurts aggregate demand and thereby lowers the return on bank assets.

The reason that, for any $\eta \geq \overline{\eta}$, the exact values of λ and η have no effect on whether deposit freezes can prevent runs is that DM output and loan defaults depend only on the total amount of cash paid out to impatient depositors. However, different values of λ and η do affect the distribution of cash among impatient depositors in a run. Keeping all else the same, a higher λ makes payouts to impatient depositors in a run more unequal since a larger share of depositors manages to redeem their deposit in full, which means less cash will be left for those impatient depositors not among the first λ to arrive. Similarly, keeping λ fixed, ex post payouts in a run become more unequal as banks increase η: while those (relatively) early in line can withdraw a larger amount, a larger share of depositors cannot withdraw anything in the DM since banks run out of cash before all depositors are served.

Assuming that preventing runs with deposit freezes is not feasible, how should banks set η in order to minimise ex post welfare losses caused by a run? Given that aggregate DM output and loan defaults are the same for all $\eta \geq \overline{\eta}$, and given that DM preferences are strictly concave, it is straightforward that the best banks can do is to distribute the cash as evenly as possible among the impatient depositors who did not manage to withdraw their deposit in full. This leads to the following proposition, which we state without separate proof:

Proposition 5. *With pure deposit freezes, the ex post welfare loss of a run is minimised if banks impose a deposit freeze with $\eta = \overline{\eta}$ once they realise that a run is underway.*

6.2. Redemption penalties

We now turn to the case where, in addition to freezing part of deposits, banks can impose penalties on DM redemptions once they realise that a run is underway. To streamline the exposition, we will take it as given that banks set $\eta \in [\overline{\eta}, \overline{\overline{\eta}}]$ in the remainder of this section.[19]

A given mix of deposit freezes and redemption penalties (η, σ) *stops* a run if a patient depositor (who is not among the first λ of depositors in the queue) is better off not withdrawing in the DM even if (hypothetically) all other depositors continue running after banks have imposed (η, σ).[20] If all depositors continue withdrawing in the DM after banks have imposed (η, σ), then we have $\omega = \overline{\omega}(\eta)$ and $\chi = m_S^b / d_S^I$ (for the latter, see the discussion in the previous subsection). From (38), we obtain that CM payouts in such a situation equal

$$d_R^P\left(\frac{m_S^b}{d_S^I}, \overline{\omega}(\eta); \eta, \sigma\right) \equiv \underline{d}_R^P(\eta, \sigma). \tag{41}$$

Banks' reaction to a run thus stops the run if and only if

$$\eta d_S^I + (1 - \eta - \sigma)\underline{d}_R^P(\eta, \sigma) \leq \underline{d}_R^P(\eta, \sigma). \tag{42}$$

Proposition 6. *There always exists a policy (η, σ) with $\eta > \overline{\eta}$ that stops a run.*

Proposition 6 states that banks can always stop a run without completely freezing deposits (which trivially stops the run). Specifically, banks can stop runs by setting η low enough and σ high enough. We prove Proposition 6 by showing in Appendix A.5.1 that condition (42) can be reformulated as a lower bound on the redemption penalty, $\sigma \geq \underline{\sigma}(\eta)$, and that there exists a value $\tilde{\eta}^{max} > \overline{\eta}$ such that $\underline{\sigma}(\eta) \in [0, 1-\eta]$ for all $\eta \in [\overline{\eta}, \tilde{\eta}^{max}]$. Therefore, any policy with $\eta \in [\overline{\eta}, \tilde{\eta}^{max}]$ and $\sigma \geq \underline{\sigma}(\eta)$ stops a run.

Banks' reaction to runs *prevents* a run if the prospect that banks will impose (η, σ) once they realise a run is underway eliminates patient depositors' incentive to withdraw their entire deposit *before* restrictions are imposed. While stopping a run ex post is necessary to prevent a run ex ante, it is not sufficient. Unlike with pure deposit freezes, stopping and preventing runs are thus not equivalent with redemption penalties.

Suppose banks' reaction to runs satisfies condition (42), such that it stops the run. Suppose also for the moment that all remaining impatient depositors withdraw in the DM after banks have imposed (η, σ). Then, we have $\omega = \theta(1 - \lambda)$ and $M^{RE} = \theta\,[\lambda + (1 - \lambda)\eta]\,d_S^I$. By Proposition 2, the share of non-defaulting entrepreneurs then equals

$$\chi = \lambda + (1 - \lambda)\eta \tag{43}$$

and, by (38), CM payouts equal

$$d_R^P\big(\lambda + (1 - \lambda)\eta, \theta(1 - \lambda); \eta, \sigma\big) \equiv \overline{d}_R^P(\eta, \sigma). \tag{44}$$

Of course, impatient depositors must be willing to withdraw in the DM and incur the redemption penalty rather than to wait until the CM, which requires

$$u(\eta q_S) + \beta\phi_+(1 - \eta - \sigma)\overline{d}_R^P(\eta, \sigma) \geq \beta\phi_+\overline{d}_R^P(\eta, \sigma), \tag{45}$$

where ϕ_+ denotes the value of money next period.[21] In Appendix A.5.2, we show that condition (45) can be rewritten as an upper bound on the redemption penalty, $\sigma \leq \overline{\sigma}(\eta)$. Finally, given that banks' reaction to runs satisfies conditions (42) and (45), patient depositors' incentive to run in the first place is eliminated if and only if

$$d_S^I \leq \overline{d}_R^P(\eta, \sigma). \tag{46}$$

In Appendix A.5.3, we show that condition (46) can be rewritten as a lower bound on the redemption penalty, $\sigma \geq \hat{\underline{\sigma}}(\eta)$. Note that the lower bounds on the redemption penalty resulting from conditions (42) and (46) are distinct. The former states that the redemption penalty must be high enough to deter patient depositors from running after the penalty has been imposed; the latter states that the redemption penalty incurred by impatient depositors must be high enough to deter patient depositors from running in the first place.

In sum, in order to prevent runs, banks' reaction to runs must satisfy conditions (42), (45) and (46).

Proposition 7. *Even with $\lambda = 0$, there may be no policy (η, σ) that prevents a run.*

We prove Proposition 7 by providing in Appendix A.7 an example where there is no (η, σ) preventing runs even if $\lambda = 0$. However, in many cases, banks can prevent runs by setting (η, σ) appropriately as long as λ is low enough; we provide an example for this case below. Note in particular that the redemption penalty σ has a redistributive function, in the sense that it redistributes funds from (impatient) depositors who

[19] If $\eta < \overline{\eta}$, banks do not pay out all their cash during a run, which, as discussed above, exacerbates the incentive to run. If $\eta > \overline{\overline{\eta}}$, banks' cash reserves are not sufficient to pay out all impatient depositors in the DM, even if patient depositors stop running after banks impose the partial freeze. Banks could then always adjust (η, σ) in such a way that CM payouts and hence run incentives are not affected but the available cash is distributed more evenly among impatient depositors in a run.

[20] Since withdrawal decisions of patient depositors are strategic complements, this is the same as saying that not withdrawing ηd_S^I in the DM must be the dominant strategy for patient depositors.

[21] Since loan defaults are decreasing in the number of impatient depositors who redeem in the DM, withdrawal decisions of impatient depositors in the DM are strategic substitutes. Condition (45) is thus the same as saying that withdrawing in the DM after banks have imposed the redemption penalty must be the dominant strategy for impatient depositors.

redeem in the DM to (patient) depositors who redeem in the CM. If (η, σ) is such that defaults caused by a run are kept sufficiently low, and the redistribution towards patient depositors implemented by the redemption penalty in the event of a run is sufficiently large, then patient depositors' incentives to participate in the run in the first place is eliminated.

6.3. Minimising welfare losses caused by runs

As with pure deposit freezes, we can ask how banks should set (η, σ) so as to minimise ex post losses caused by runs, assuming preventing runs is not feasible. Note first that if banks fail to stop the run, then the situation is equivalent to the one with pure deposit freezes: as long as $\eta \geq \overline{\eta}$, the total cash paid out to impatient depositors in a run equals θm_S^b, which in turn pins down DM output and loan defaults. However, with redemption penalties, banks can do better by deterring patient depositors from withdrawing once the run has been detected, which allows to increase the aggregate amount of cash paid out to impatient depositors.

Given $\eta \leq \overline{\overline{\eta}}$, all impatient depositors who are not among the first λ of depositors to arrive in a run receive identical DM payouts whenever (η, σ) is such that the run is stopped once it is detected. Minimising welfare losses caused by a run is then equivalent to maximising DM output in a run, which is achieved by maximising the cash paid out to impatient depositors (i.e. maximising η) subject to the relevant constraints (42) and (45). Specifically, DM output in a run is maximised by setting η to the highest level within $(\overline{\eta}, \overline{\overline{\eta}}]$ consistent with stopping the run while at the same time ensuring that impatient depositors are willing to withdraw. We denote this level with $\eta^{\max}$.

Proposition 8. *Suppose $\lambda < \theta$. Let $\eta^{\max} \equiv \min\{\overline{\overline{\eta}}, \tilde{\eta}^{max}, \hat{\eta}^{max}\}$, where $\tilde{\eta}^{max}$ is the unique value of η solving $\underline{\sigma}(\eta) = 1 - \eta$, and $\hat{\eta}^{max}$ is the unique strictly positive value of η solving $\underline{\sigma}(\eta) = \overline{\sigma}(\eta)$. We have $\eta^{max} \in (\overline{\eta}, 1)$, and the ex post welfare loss of a run is minimised if banks set $\eta = \eta^{\max}$ and $\sigma \in [\underline{\sigma}(\eta^{\max}), \overline{\sigma}(\eta^{\max})]$.*

We refer to Appendix A.6 for the proof and the derivations related to Proposition 8. Intuitively, the maximum amount of money that can be paid out to impatient depositors in the DM in case of a run can be constrained for three different reasons. It can be constrained by the fact that banks have limited cash left once they notice that a run is underway (in which case $\eta^{\max} = \overline{\overline{\eta}}$), it can be constrained by the fact that patient depositors must be deterred from continuing running while the redemption penalty cannot exceed the fraction of frozen deposits (in which case $\eta^{\max} = \tilde{\eta}^{max}$), or it can be constrained because patient depositors must be deterred from continuing running while impatient depositors must still find it attractive to withdraw (in which case $\eta^{\max} = \hat{\eta}^{max}$).

6.4. Numerical example (where runs can be prevented)

Fig. 1 provides a numerical example illustrating how (η, σ) needs be set in order to both stop and prevent runs. We assume that the DM utility function equals $u(q) = q^{\nu}$ and that DM production follows a Cobb–Douglas technology, i.e.

$$f(k, h) = k^{\alpha} h^{1-\alpha}. \tag{47}$$

The parameters used for this example are summarised in Table 1; they are such that the economy is in a ZLB equilibrium. The grey area in Fig. 1 depicts the set of (σ, η) that both stop and prevent runs. The fact that $\underline{\sigma}(\eta)$ is strictly increasing in η while $\underline{\hat{\sigma}}(\eta)$ is strictly decreasing in η is not specific to the example at hand; as we show in Appendix A.5, this is always the case if banks are fragile under pure deposit freezes. Higher DM payouts η increase the incentive for patient depositors to continue running, which is why the redemption penalty required to stop the run increases in η. Furthermore, given that the redemption penalty stops runs, higher DM payouts to impatient depositors lead to more DM activity and less defaults, thereby decreasing patient depositors' incentives to participate in the run in the first place. For this reason, the minimum redemption penalty necessary to prevent a run is decreasing in η. This is a characteristic property of our model: increasing payouts to impatient depositors can lower patient depositors' incentives to run due to the tight connection between banks' liquidity provision and the return on bank assets.

Table 1
Parameter values for Fig. 1.

α	ν	θ	n	λ	ι	B	q_S	ρ_S	κ_S
0.6	0.65	0.6	0.6	0.45	0.05	0	0.037	1.904	1.199

7. Emergency liquidity

We now turn to the role of policy in stopping and preventing runs. For expositional convenience, we assume the government provides emergency liquidity by purchasing assets from banks. We show in Appendix A.13 how the intervention can be reinterpreted as the provision of collateralised loans to banks without changing any of the results. We refer to Appendices A.8–A.12 for the proofs in this section that are not contained in the main text.

To provide emergency liquidity, the government stands ready to convert bonds and loans with a gross face value (i.e. principal plus interest) of one dollar into δ_b and δ_ℓ dollars of cash, respectively, where we assume $\delta_\ell \leq \delta_b \leq 1$. Here, $\delta_b < 1$ and $\delta_\ell < 1$ can be interpreted as discounts on bond and loan purchases, with lower values of δ_b and δ_ℓ representing higher discounts. For simplicity, we assume throughout this section that $\lambda = \theta$, i.e. the government provides emergency liquidity after a measure θ of depositors has withdrawn.

We allow for the possibility that the government is less efficient than banks in collecting loan repayments from entrepreneurs. Specifically, when the government purchases an amount ℓ of loans and the default rate is $1 - \chi$, the government earns $\zeta \chi (1 + i_S)\ell$ from its loan portfolio, where $\zeta \in (0, 1]$ captures the government's efficiency relative to banks in collecting loan repayments.

Furthermore, we assume that the government cannot increase its real indebtedness, which restricts the ability to provide liquidity in real terms.[22] Specifically, we have

$$\frac{M + (1 + i_S)B}{M_S + (1 + i_S)B_S} \leq \frac{1 + \pi}{1 + \pi_S}, \tag{48}$$

where $M + (1 + i_S)B$ denotes the government's nominal liabilities at the beginning of the next CM after entrepreneurs have repaid loans, and $1 + \pi = \phi_t / \phi_{t+1}$ is the inflation rate given the government's provision of emergency liquidity. Condition (48) implies that the provision of emergency liquidity will be inflationary whenever it leads to an increase in nominal government liabilities beyond the point at which the assets purchased by the government mature.

Consider now the provision of emergency liquidity from the point of view of banks. With emergency liquidity, the total amount of money banks can access once the run is detected is

$$m_S^b - \theta d_S^I + \delta_b (1 + i_S) b_S^b + \delta_\ell (1 + i_S) \ell_S^b. \tag{49}$$

Following the results in Section 6, we assume that once banks notice a run is underway, they charge a redemption penalty that reflects the discounts imposed on asset purchases and thus ensures that withdrawing

[22] This constraint can be motivated in various ways. For instance, it may capture the central bank's inability to raise taxes, or it may capture political constraints such as a limit on real government debt.

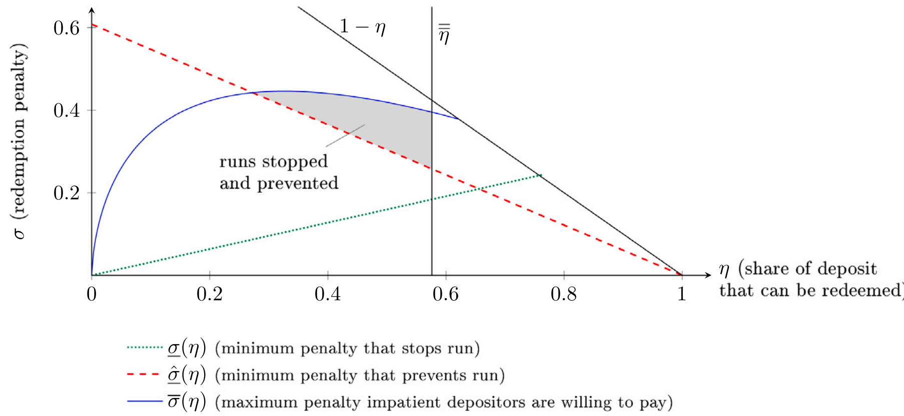

Fig. 1. Example where runs can be prevented.

depositors internalise the cost that banks incur when selling assets:[23]

$$\eta = 1 - \sigma = \min\left\{ \frac{m_S^b - \theta d_S^I + \delta_b(1+i_S)b_S^b + \delta_\ell(1+i_S)\ell_S^b}{(1-\theta)d_S^I}, 1 \right\}. \tag{50}$$

Lemma 1. *In ZLB and PLI equilibria, banks charge a redemption penalty if and only if $\delta_\ell < 1$, i.e. if and only if the government imposes a discount on its asset purchases. In FLI equilibria, banks charge a redemption penalty if and only if discounts are sufficiently high.*

Proof. From (50), we get

$$\eta < 1 \;\Leftrightarrow\; m_S^b + (1+i_S)(\delta_\ell \ell_S^b + \delta_b b_S^b) < d_S^I. \tag{51}$$

Recall that $d_S^I = m_S^b + (1+i_S)(\ell_S^b + b_S^b)$ in ZLB and PLI equilibria and $d_S^I \le m_S^b + (1+i_S)(\ell_S^b + b_S^b)$ in FLI equilibria, with strict inequality except for a knife-edge case. Together with condition (51) and the fact that $\delta_b \ge \delta_\ell$, the result in Lemma 1 follows. ■

The redemption penalty in (50) ensures that even if patient depositors continue withdrawing after the penalty is imposed, banks can pay out all remaining $1-\theta$ depositors without running out of funds. Therefore, regardless of whether the run stops once the penalty is imposed, the aggregate amount of cash paid out to impatient depositors in a run equals[24]

$$M^{RE} = \theta\,[\theta + \eta(1-\theta)]\,d_S^I. \tag{52}$$

Next, we denote τ_b and τ_ℓ as the share of banks' bond and loan holdings sold to the government, and we let $\mathcal{I} \in \{0,1\}$ be an indicator variable that equals one if the run continues after a measure θ of depositors has withdrawn. Banks' asset sales are such that

$$\eta(1-\theta)\,[\theta + \mathcal{I}(1-\theta)]\,d_S^I = (m_S^b - \theta d_S^I) + \tau_b\delta_b(1+i_S)b_S^b + \tau_\ell\delta_\ell(1+i_S)\ell_S^b, \tag{53}$$

where the LHS equals the money needed to meet additional withdrawals once the run has been detected, and the RHS equals excess reserves plus the money obtained from selling assets to the government. In the following, we assume without loss of generality that $\tau_b = \tau_\ell = \theta + (1-\theta)\mathcal{I}$, i.e. the fraction of bonds and loans sold to the government equals the fraction of remaining depositors that continue withdrawing once the run is detected. Banks' CM payout to non-withdrawing depositors after a run in the DM is then given by

$$\frac{m_S^b - \theta d_S^I + (1+i_S)(b_S^b + \chi \ell_S^b)}{1-\theta}. \tag{54}$$

Using (54), we can now derive the conditions under which emergency liquidity (i) stops a run once it has been detected and (ii) prevents runs from starting in the first place. A run stops once it has been detected if and only if

$$\eta d_S^I \le \frac{m_S^b - \theta d_S^I + (1+i_S)(b_S^b + \chi \ell_S^b)}{1-\theta}, \tag{55}$$

i.e. if and only if patient depositors are better off not withdrawing in the DM once the redemption penalty is imposed. Next, the provision of emergency liquidity prevents runs from starting in the first place if and only if

$$d_S^I \le \frac{m_S^b - \theta d_S^I + (1+i_S)(b_S^b + \chi \ell_S^b)}{1-\theta}, \tag{56}$$

i.e. if and only if patient depositors are better off not withdrawing their entire deposit in the DM before a redemption penalty is imposed, given that emergency liquidity will be provided in a run. As is clear from condition (56), whether emergency liquidity prevents runs boils down to whether it prevents a (too large) drop in loan repayments, χ, in case of a run.

Lemma 2. *In ZLB and PLI equilibria, emergency liquidity prevents runs if and only if $\chi = 1$, i.e. if and only if loan defaults in case of a run are avoided completely. In FLI equilibria, emergency liquidity prevents runs if and only if χ is sufficiently high.*

Proof. The result in Lemma 2 follows directly from condition (56) together with the fact that in ZLB and PLI equilibria, we have $d_S^I = m_S^b + (1+i_S)(\ell_S^b + b_S^b)$, while in FLI equilibria, we have $d_S^I \le m_S^b + (1+i_S)(\ell_S^b + b_S^b)$, with strict inequality except for a knife-edge case. ■

The degree to which emergency liquidity stabilises loan repayments in a run depends on the degree to which it stabilises impatient depositors' purchasing power in the DM relative to the steady state. As is rather obvious, we get from (50) and (52) that the government can always prevent a fall in impatient depositors' *nominal* purchasing power by buying assets from banks at face value (i.e. $\delta_b = \delta_\ell = 1$), in which case banks can repay all outstanding deposits in a run without charging a redemption penalty. However, since the injection of emergency liquidity may be inflationary, this does not guarantee that impatient depositors' *real* purchasing power is maintained. A fall in impatient depositors' real purchasing power, in turn, can mean that some entrepreneurs prefer not to work in the DM and default on their

[23] The penalty in (50) is the lowest one that may potentially stop the run. If banks set a lower σ (i.e. a higher η), then a run cannot be stopped once it has started since the depositors at the end of the queue receive nothing in the DM and also nothing in the CM (the bank has sold all its assets before all depositors are served).

[24] We limit our attention to cases where the discounts, and hence the redemption penalty, are sufficiently small that impatient depositors prefer to withdraw in the DM even if they are subject to the penalty.

loans. The effect of emergency liquidity on inflation is thus key, and we will turn to this question next.

Emergency liquidity and inflation. To find out how emergency liquidity affects inflation, we need to determine its effect on the government's total nominal liabilities. At the beginning of the CM following the intervention in the DM, after entrepreneurs have repaid their loans, the government's outstanding nominal liabilities equal

$$M_S + [\tau_b \delta_b (1+i_S) b_S^b + \tau_\ell \delta_\ell (1+i_S) \ell_S^b] - \tau_\ell \zeta \chi (1+i_S) \ell_S^b$$
$$+(1-\tau_b)(1+i_S) b_S^b + (1+i_S)(B_S - b_S^b). \tag{57}$$

The first two terms in (57) are the total money in circulation after the injection of emergency liquidity in the DM, the third term is the money that the government receives back in the beginning of the CM from maturing loans purchased in the DM, and the last two terms are government bonds held by banks and households, respectively. Combining (57) with (48), we obtain[25]

$$\frac{1+\pi}{1+\pi_S} = \max\left\{1 + \frac{(1+i_S)(\delta_\ell - \zeta\chi)\tau_\ell \ell_S^b - (1+i_S)(1-\delta_b)\tau_b b_S^b}{M_S + (1+i_S)B_S}, 1\right\}. \tag{58}$$

It is immediate from (58) that the provision of emergency liquidity can only be inflationary if $\zeta\chi < \delta_\ell$, i.e. if the money that the government receives back from maturing loans is less than what it paid for these loans. Put differently, emergency liquidity can only be inflationary if the government makes a loss with its intervention. If the intervention is not loss-making, then the amount of money withdrawn from circulation when assets mature is at least as large as the amount injected in the run, and the provision of emergency liquidity does not lead to a lasting increase in nominal government liabilities.

A result that will be useful in the following analysis is that in equilibrium, an increase in inflation – with its associated decrease in impatient depositors' real purchasing power – goes in hand with a decrease in the real price of the DM good:

Lemma 3. *The real DM price, ρ, is a strictly decreasing function of π.*

The next result states that there is a critical threshold for δ_ℓ above which the government makes a loss with its loan purchases, and the threshold is increasing in the government's efficiency in collecting loan repayments:

Proposition 9. *There exists a threshold $\bar{\delta}_\ell(\zeta) \in (0,1]$ such that $\zeta\chi < \delta_\ell$ if and only if $\delta_\ell > \bar{\delta}_\ell(\zeta)$. The threshold $\bar{\delta}_\ell(\zeta)$ is strictly increasing in ζ, with $\bar{\delta}_\ell(1) = 1$. If $\delta_\ell \le \bar{\delta}_\ell(\zeta)$, then $\pi = \pi_S$, i.e. emergency liquidity is non-inflationary.*

One implication of Proposition 9 is that the government never makes a loss with its intervention if it is as efficient as banks in collecting loan repayments. The result that the provision of emergency liquidity is never loss-making when δ_ℓ is set sufficiently low is not obvious a priori. While reducing δ_ℓ means that the government buys loans at a lower price, it can also mean that banks need to impose a higher redemption penalty, which causes a fall in aggregate DM spending and thereby leads to more defaults on the loans purchased by the government. However, note that in the steady state, only a fraction of entrepreneurs' DM revenue goes towards loan repayments since entrepreneurs keep a share of their revenue as compensation for their labour effort. Therefore, as long as the share of entrepreneurs' revenue going to loan repayments remains constant (which is the case when inflation remains constant), a given decrease in aggregate DM spending translates into a smaller decrease in aggregate loan repayments. As a result, setting δ_ℓ below $\bar{\delta}_\ell(\zeta)$ never leads to net losses for the government.

Non-inflationary emergency liquidity. We now consider the case where the government sets $\delta_\ell \le \bar{\delta}_\ell(\zeta)$, which means the provision of emergency liquidity is not inflationary. Whether non-inflationary emergency liquidity stops and prevents runs depends on its effect on loan repayments in a run, χ. From (52) together with our result in Proposition 2, we have that with non-inflationary emergency liquidity,

$$\chi = \frac{M^{RE}}{\theta d_S^I} = \theta + \eta(1-\theta), \tag{59}$$

i.e. the higher the penalty on DM redemptions, the more defaults occur. Defaults are avoided only if banks do not charge a penalty ($\eta = 1$). The following result shows that runs can always be stopped with the provision of non-inflationary emergency liquidity:

Proposition 10. *Emergency liquidity with $\delta_\ell \le \bar{\delta}_\ell(\zeta)$ stops runs.*

Proof. Suppose first that the discounts $\delta_\ell \le \bar{\delta}_\ell(\zeta)$ and δ_b are such that banks do not need to impose a redemption penalty, i.e. $\eta = 1$. By (59), this implies $\chi = 1$, in which case condition (55) is evidently fulfilled since $d_S^I \le m_S^b + (1+i_S)(\ell_S^b + b_S^b)$. Suppose next that the discounts are such that banks need to impose a redemption penalty, i.e. $\eta < 1$. Substituting for η using (50), condition (55) then becomes $\delta_b b_S^b + \delta_\ell \ell_S^b \le b_S^b + \chi \ell_S^b$. Since $\delta_b \le 1$, a sufficient condition that runs are stopped is that $\chi \ge \delta_\ell$, which, by Proposition 9, is the case when $\delta_\ell \le \bar{\delta}_\ell(\zeta)$. ■

The result in Proposition 10 is closely related to the fact that, as already discussed above, a given decrease in DM spending translates into a smaller decrease in loan repayments when inflation remains constant. Therefore, even if the discounts are such that banks need to impose a redemption penalty – which leaves impatient depositors with less cash to spend and thereby causes a certain amount of loan defaults – the losses on the loans held by banks are always smaller than the redemption penalty. This in turn means that patient depositors are better off withdrawing in the CM than paying the redemption penalty in the DM.

Let us now turn to the question whether non-inflationary emergency liquidity can prevent runs from starting in the first place. From (59), we know that defaults in a run increase in the redemption penalty, which in turn increases in the discounts charged by the government on its asset purchases. Therefore, to minimise the incentive to run with non-inflationary emergency liquidity, discounts should be set to the lowest levels consistent with preventing losses, i.e. $\delta_\ell = \bar{\delta}_\ell(\zeta)$ and $\delta_b = 1$. This leads to the following result:

Proposition 11. *In ZLB and PLI equilibria, non-inflationary emergency liquidity can prevent runs if and only if $\zeta = 1$. In FLI equilibria with $\mathcal{A}_S \le \bar{\mathcal{A}}$, non-inflationary emergency liquidity can prevent runs if and only if,*[26]

$$\zeta \ge \underline{\zeta}, \quad \text{where } \underline{\zeta} \text{ solves} \quad \bar{\delta}_\ell(\underline{\zeta}) = 1 - \frac{\theta\kappa^* + \rho_S}{(\theta\kappa^*)^2 q_S}\left[\mathcal{A}_S - (1-\theta)\rho_S q_S\right]. \tag{60}$$

Let us briefly elaborate on the result in Proposition 11. Consider first ZLB and PLI equilibria, for which we know from Lemma 2 that runs are only prevented if emergency liquidity avoids any loan defaults. We also know from Lemma 1 that in ZLB and PLI equilibria, any strictly

[25] In case the government makes a profit with its intervention, we assume it puts these profits back into circulation as money, which prevents the intervention from being deflationary.

[26] Recall that if $\mathcal{A}_S > \bar{\mathcal{A}}$, not all assets in the economy are held by banks. As in *Proposition* 4 whether banks are fragile then depends additionally on how illiquid assets are allocated between households and banks; for a given ζ, banks are more (less) likely to be fragile if they hold a large share of outstanding loans (bonds).

positive discounts on asset purchases force banks to charge a redemption penalty, which, by (59), means that there will be defaults. As a result, non-inflationary liquidity can only prevent runs in ZLB in PLI equilibria if the government purchases assets at face value ($\delta_\ell = \delta_b = 1$). From (58), we know that this is only non-inflationary if the government makes no losses on its loan purchases, which, by Proposition 9, requires that the government be as efficient as banks in collecting loan repayments ($\zeta = 1$). Consider next FLI equilibria, for which Proposition 11 states that runs can be prevented with non-inflationary emergency liquidity as long as the government is sufficiently efficient in collecting loan repayments.[27] In FLI equilibria, banks have a buffer allowing them to absorb some discounts on asset purchases without having to charge a redemption penalty, and also to absorb some loan defaults without becoming susceptible to runs. It is therefore possible for the government to impose strictly positive discounts on loan purchases in FLI equilibria without making the banking system fragile, as long as the discounts are not too high.

Inflationary emergency liquidity. We now turn to the case where the government sets $\delta_\ell > \overline{\delta}_\ell(\zeta)$, which means the provision of emergency liquidity can be inflationary. In particular, the question arises whether inflationary emergency liquidity can prevent runs in cases where non-inflationary emergency liquidity fails to do so due to the required discount on loan purchases.

Recall that whether emergency liquidity prevents runs comes down to whether it averts a (too large) fall in loan repayments, χ, in case of a run. From (31), we know that χ depends on (i) the total nominal spending of impatient depositors, M^{RE}, and (ii) the share of entrepreneurs' revenue available for loan repayments, α, which in turn is a function of the real DM price, ρ. Since emergency liquidity can always stabilise nominal spending of impatient depositors by setting $\delta_b = \delta_\ell = 1$, the key question is how inflation affects α. From Lemma 3, we know that an increase in inflation implies a fall in ρ. We then obtain the following result:

Proposition 12. *If the capital share $\alpha(\rho)$ is non-increasing in ρ, then emergency liquidity with $\delta_\ell = \delta_b = 1$ ensures $\chi = 1$ and thus prevents runs. If $\alpha(\rho)$ is strictly increasing in ρ, then emergency liquidity cannot prevent runs if $\zeta < 1$ and the economy is in a ZLB or PLI equilibrium.*

Proposition 12 tells us that stabilising *nominal* DM demand is sufficient to eliminate loan defaults as long as an increase in inflation – with the attendant decrease in ρ – does not lead to a fall in the capital share.[28] What is key for this result is that inflation is a double-edged sword for entrepreneurs: on the one hand, inflation reduces impatient depositors' real purchasing power, but on the other hand, it also reduces entrepreneurs' real debt burden. Since these two effects are proportional, inflation does not negatively affect entrepreneurs' incentives to produce in the DM and repay debt as long as it does not lead to a decrease in α. Conversely, if inflation does lead to a decrease in α, then inflation inevitably goes in hand with some loan defaults; and if banks have no buffer to absorb such losses, as is the case in ZLB and PLI equilibria, inflationary emergency liquidity cannot prevent runs.[29]

To see how α moves with changes in ρ (and hence with changes in π), recall first from Section 5 that a fall in ρ implies a fall in DM output per active entrepreneur, q^e. Since entrepreneurs' capital stock is fixed at k_S, any change in q^e is entirely due to a change in entrepreneurs' hours worked, h. How α changes with ρ thus depends on how α changes when entrepreneurs alter production by varying h given k_S. This in turn depends on the properties of the DM production function; with a Cobb–Douglas production function, for instance, α is constant.[30]

Real loan contracts. To derive the above result that inflationary emergency liquidity can eliminate runs if $\alpha(\rho)$ is weakly decreasing, it was important that inflation reduces entrepreneurs' real debt burden. This raises the question whether inflationary emergency liquidity can still prevent runs if loan contracts are denominated in real terms, in which case the gross nominal repayment on a loan ℓ equals $\frac{1+\pi}{1+\pi_S}(1+i_S)\ell$. Somewhat counterintuitively, inflationary emergency liquidity is *more* effective in preventing runs when loan contracts are real, in the sense that stabilising nominal DM demand is sufficient to eliminate runs regardless of the details of the DM production function:

Proposition 13. *If loan contracts are denominated in real terms, emergency liquidity with $\delta_\ell = \delta_b = 1$ prevents runs.*

To understand this surprising result, recall that (as we showed in the proof of Proposition 2) the equilibrium real DM price cannot deviate from its steady state value, ρ_S, when entrepreneurs' real debt burden is fixed: $\rho < \rho_S$ would imply that entrepreneurs' payoff from producing any amount $q^e > 0$ is negative, while $\rho > \rho_S$ is not possible either, since this would imply higher DM production than in the steady state, which violates DM market clearing given that impatient depositors' real purchasing power cannot increase in a run. Therefore, with real loan contracts, any fall in DM output caused by inflation occurs entirely on the extensive margin, while production of each non-defaulting entrepreneur remains the same as in the steady state. This means that inflation does not affect the capital share $\alpha(\rho)$, which in turn implies that banks' total revenue on their loan portfolio does not suffer as long as policy stabilises aggregate nominal DM spending; any decrease in nominal repayments by non-producing entrepreneurs is exactly offset by an increase in nominal repayments of producing entrepreneurs. The fact that policy prevents losses on banks' loan portfolios then removes any incentive to run on the banks in the first place.

8. Empirical relevance

Even though we view our paper primarily as a theoretical contribution, we also want to discuss briefly how we interpret the relevance of our mechanism in the empirical context. In our model, depositors withdraw cash from banks in a run, which hinders economic activity in sectors relying on cash transactions. Therefore, our model may be particularly relevant to cash intensive economies. In line with that, we think there is ample anecdotal evidence for the mechanism highlighted in our paper during the three major financial crises in the U.S. National Banking Era (1864-1913) in 1873, 1893 and 1907.[31]

During each of these episodes, New York City banks suspended convertibility of deposits into cash, which quickly led to suspension throughout the country. In all three crises, the suspension lasted for several weeks and in the case of 1907 for two months. Contemporary reports on the negative effect of suspension on real economic activity abound. In his classic study on the National Banking Era crises, Sprague (1910) lists for 1873 numerous excerpts from local newspapers reporting that firms had to temporarily curtail production after suspension

[27] From (28), we have $\mathcal{A}_S > (1-\theta)\rho_S q_S$ and hence $\underline{\zeta} < 1$ in FLI equilibria except for a knife-edge case where $\underline{\zeta} = 1$.

[28] The notion that preventing runs may require the central bank to commit to an inflationary policy in the event of a run (off the equilibrium path) is also found in Schilling et al. (2024). They highlight that such policies may not be credible if increases in inflation negatively enter the central bank's objective function.

[29] In FLI equilibria, banks have some buffer to absorb losses on their loans; runs may then be prevented with inflationary emergency liquidity even if inflation leads to a decrease in α, as long as the decrease is not too large.

[30] More generally, it can be shown that if $f(k,h)$ is a constant elasticity of substitution (CES) production function with elasticity of substitution ϵ, then $\alpha'(\rho) > 0$ if and only if $\epsilon < 1$.

[31] Although many transactions could be settled with bank checks and drafts at the time, cash was dominant for wage payments and retail transactions (James et al., 2013).

was imposed, in large part due to their inability of obtaining cash for wage payments (pp. 71–74).[32] For 1893 as well, Sprague (1910, pp. 199–203) describes the deleterious effects of suspension, including a steep drop in railway earnings after suspension was imposed (which Sprague attributes in large part to the paralysing effects of the cash shortage) and newspaper reports of temporary factory closures due to an inability to make collections and obtain cash to meet payrolls. Sprague (1910) concludes that 'There is [..] evidence that suspension was a potent factor accentuating the depression in trade in the month of August [1893].' (p.200) and reports a rapid business recovery once the restrictions were lifted, which in his view '[..] affords striking evidence of the disturbing effect which had been brought about by suspension.' (p.203). Similar accounts can be found for 1907. For instance, the Comptroller of the Currency (1907) states with regard to the 1907 suspension that 'The result was to at once precipitate a most serious bank crisis and a famine of currency for pay rolls and other necessary cash transactions.' (p.70) and notes further that 'The greatest hardship to business generally has been the derangement of the machinery for making collections and remittances. As can readily be seen, this has interfered with every kind and class of business and led to great curtailment of business operations of every kind. Factories have suspended, workmen have been thrown out of employment, orders have been cancelled, the moving of crops has been greatly retarded and interfered with and exports have fallen off at a time of the year when they should be at their highest.' (pp. 70–71).[33]

Furthermore, Sprague (1910) provides evidence that many individuals and businesses stored money in safes or deposit boxes during the National Banking Era panics. The Comptroller of the Currency (1907) states with regard to the 1907 crisis that 'Money has been withdrawn and hoarded by individuals, corporations, and [..] by the banks themselves, all of whom at once drew and held all the money of any kind they could obtain, often really in larger sums than needed.' (p.70).[34] These accounts of currency hoarding on the one hand and of widespread cash shortages on the other hand are consistent with the mechanism in our model, in which a run leads patient depositors to withdraw money and store it, while impatient depositors curtail their consumption due to the difficulty of obtaining cash from banks.

Our model predicts that this misallocation of liquidity has negative effects on output, which in turn leads to losses on bank assets. As described in Friedman and Schwartz (1963, pp. 163–164) and James et al. (2013), contemporary observers during the National Banking Era overall had a very negative view of restrictions on cash withdrawals and often considered this measure to make the situation worse rather than better. Nevertheless, how severe the effect of suspension on economic activity was is difficult to assess due to a lack of relevant high-frequency data (see, e.g. Wicker (2000) on this point). To the best of our knowledge, the most ambitious attempt to estimate the effect of the restriction on cash payouts on economic activity during the National Banking Era crises was undertaken by James et al. (2013). They show that various proxies for economic activity display a clear drop after cash withdrawals were restricted, and once restrictions were lifted, they indicate a recovery or at least stabilisation. The econometric analysis in James et al. (2013) finds an independent negative effect of suspension on economic activity, consistent with the qualitative accounts of earlier writers. Finally, regarding the effect of suspension of deposit convertibility on bank health, Wicker (2000, Chapter 4) documents a high incidence of bank failures in 1893 even after convertibility was suspended, which he contrasts (on p.82) with Friedman and Schwartz's view that suspension of cash withdrawals should put a halt to bank failures. Jaremski and Wheelock (2023) document that the 1907 panic with its prolonged period of suspension was followed – with some delay – by a significant increase in bank closures. Overall, the historical discussion and more recent analysis of these episodes indicate that the suspensions adversely affected economic activity and correspondingly bank health, precisely as in our model.

We believe that our model can be relevant in a modern context as well for several reasons. First, even in modern times there are economies, including some of the larger emerging markets, that still operate to a significant degree on a cash basis (Bech et al., 2018; World Bank, 2021). A relatively recent example is Argentina, where a partial suspension of cash withdrawals from banks (dubbed the *corralito*) was introduced in 2001 in response to a countrywide bank run.[35] Many observers considered the suspension to be highly disruptive to the Argentinian economy. For instance, Kiguel (2011) states that 'Under this system [of partial suspension], people could only transfer funds within the banking system but they were not allowed to get cash, except in small amounts. This measure resulted in a monetary crunch and led to a collapse of economic activity – especially in the informal sector which mainly works on cash [..]' (p.7). Similarly, Saxton (2003) states with reference to the Argentinian deposit freeze that 'The economy turned from recession to depression as people and businesses could not make payments.' (p.12). Daseking et al. (2004) document a veritable collapse of economic activity in the month after the freeze was imposed and (together with some other measures taken) consider the *corralito* to 'have exacerbated the macroeconomic consequences of the crisis and complicated its resolution.' (p. 42). The Argentinian case suggests that even in a more modern context, restricting deposit convertibility can disrupt economic activity in a similar way to how we model it in our paper. It is not too far-fetched to assume that this has negative ramifications for bank assets, in which case the main ingredients of our mechanism would be present. Indeed, Gutierrez and Montes-Negret (2004) connect the negative effects of the *corralito* directly to bank health, writing: '[..] the deposit freeze not only jammed the payment system, but also destroyed the chain of payments in the economy, leading to an additional deterioration of the banks' loan portfolios.' (p.16).

Even in present-day economies in which the financial infrastructure allows almost all payments to be made directly via deposit transfers without going through cash, a version of our mechanism might still apply. In principle, one might think that a suspension of deposit convertibility into cash should have little effect on economic activity in such a financial system as long as deposit transfers are still possible. However, this is based on the assumption that sellers accept payments in deposits whose convertibility is suspended; and in a situation in which the banking system's health is in question, this may not be the case. If sufficiently many sellers refuse to be paid in deposits, then the misallocation of cash caused by a bank run may lead to a drop in economic activity, which in turn leads to losses on bank loans, justifying both the run by patient depositors and sellers' reluctance to accept payment in non-convertible deposits. While a full analysis of this case is beyond the scope of this paper, this shows how our mechanism

[32] To give just one example, the October 13, 1873 issue of the *New York Tribune* reports that 'A number of Frankford cotton mills are running on half time in consequence of falling off of orders and difficulty of procuring currency to pay wages.'

[33] Similarly, in their account of the 1907 crisis, Friedman and Schwartz (1963, p.156) write: 'Simultaneously with this suspension, the contraction became more severe. Production, freight car loadings, bank clearings and the like all declined sharply and the liabilities of commercial failures increased sharply. Restriction of payments by banks was lifted in early 1908, and a few months later recovery got underway.' Despite these negative effects, Friedman and Schwartz famously argued that suspension of cash withdrawals was *less bad* than the large scale asset liquidations undertaken by banks during the Great Depression (which is not inconsistent with our model).

[34] 'Hoarding' by banks refers to the fact that many (country) banks increased their excess reserves during the panic. For a detailed account of currency hoarding by the public during the 1907 panic, see Andrew (1908).

[35] For a detailed description of the Argentinian deposit freeze, see Gutierrez and Montes-Negret (2004).

could also play out in a setting where payments via deposit transfers are feasible.

An additional reason why we think our model could be relevant in the future is Central Bank Digital Currency (CBDC), the introduction of which is currently pondered by many of the leading central banks. If CBDC replaces other (electronic) payment methods and becomes the sole widely accepted final means of payment, and households routinely convert bank deposits into CBDC to make payments, then one can simply reinterpret the money in our model as CBDC. And if CBDC coexists with other payment methods that allow transactions to be settled directly via deposit transfers, then a self-fulfilling process like the one described above, where doubts about bank solvency lead sellers to insist on payment in CBDC, may still be possible; indeed, one may speculate that CBDC makes this somewhat more likely due to low transaction costs for sellers of switching to a CBDC-only policy in a crisis.

Finally, we also want to highlight situations where our mechanism may not apply, even if we take it as given that most transactions are settled with cash. First, in our model, it is crucial that bank depositors and the receivers of bank loans operate in the same real markets; only in this case can restrictions on deposit withdrawals have a significant negative effect on the revenues of bank borrowers. Hence, if there is a disconnect between the origin of depositors and bank assets, our mechanism becomes less relevant, e.g. if banks have an international depositor base but invest primarily in the domestic economy or vice versa.[36] Second, our model is arguably most relevant for runs that are truly systemic in nature. When only a subset of banks is hit by a run, and both depositors and borrowers of the affected banks constitute a relatively small part of the economy, then it is less likely that the negative repercussions of withdrawal restrictions imposed by the affected banks on the revenues of borrowers from the same banks are strong enough to trigger the self-fulfilling process described in our model.

9. Conclusion

In this paper, we have studied systemic bank runs in a general equilibrium banking model in which the return on bank assets depends on banks' liquidity provision to households. In our setup, households redeem bank deposits to purchase goods from entrepreneurs, which in turn obtain loans from banks. A run leads to a misallocation of liquidity in the sense that money is withdrawn by patient households who store it, which means less money is available for impatient households to buy goods. The fact that impatient households have less money to spend then reduces demand in the goods market, which causes some entrepreneurs to default on their bank loans, rationalising the run by patient households. A key insight is that deposit freezes do not protect the value of bank assets in a run: a freeze reduces consumption demand by impatient households, leading to an economic downturn and losses on bank assets. For this reason, deposit freezes may not be a suitable means to eliminate run equilibria, even in the absence of aggregate uncertainty and commitment problems.

While our model highlights the limitations of redemption restrictions as tools to prevent runs, we have also shown that in some cases, run equilibria can be eliminated by properly calibrated redemption penalties combined with partial deposit freezes. Intuitively, the difficulty when calibrating these measures is that on the one hand, they need to be strict enough to deter patient households from withdrawing, but on the other hand, they should allow sufficient withdrawals by impatient households so that economic activity does not fall too sharply. Furthermore, we have shown that government-provided emergency liquidity can eliminate runs if it averts a fall in the real purchasing power of impatient households in a run. Even if the intervention is inflationary, it can in some cases still prevent runs, depending on how inflation affects entrepreneurs' incentives to produce and repay their bank loans.

[36] A possible example of this is Cyprus, where a deposit freeze was imposed in 2013. At the time, a large part of deposits in Cypriot banks were held by non-eurozone residents, while bank borrowers were predominantly domestic and Greek residents (European Commission, 2013).

CRediT authorship contribution statement

Lukas Altermatt: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Hugo van Buggenum:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Lukas Voellmy:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Replication Package: Systemic Bank Runs without Aggregate Risk: How a Misallocation of Liquidity May Trigger a Solvency Crisis (Original data) (Mendeley Data)

Appendix. Additional derivations and proofs

A.1. Deriving households' problem in the economy without banks

Denoting $V_t(m, \ell, b)$ as the households' CM value function in period t, and taking it as given that $p_t q_t = m_t$, we have:

$$V_t(m,\ell,b) = \max_{\{x_t^h, l_t, m_t, \ell_t, b_t\}} \Big\{ U(x_t^h) - l_t + \theta\left[u(m_t/p_t) + \beta V_{t+1}(0,\ell_t,b_t)\right] + (1-\theta)\beta V_{t+1}(m_t,\ell_t,b_t) \Big\}$$

$$\text{s.t.} \quad l_t = x_t^h + \Delta_t + \phi_t\left[(m_t - m) + (\ell_t - (1+i_{\ell,t})\ell) + (b_t - (1+i_{b,t})b)\right].$$

Inserting the households' flow budget constraint into the objective function, we can rewrite the value function as:

$$V_t(m,\ell,b) = \phi_t\left[m + (1+i_{\ell,t})\ell + (1+i_{b,t})b\right] - \Delta_t + \max_{\{x_t^h, m_t, \ell_t, b_t\}} \Big\{ U(x_t^h) - x_t^h - \phi_t(m_t + \ell_t + b_t) + \theta u(m_t/p_t) + \beta\left[\theta V_{t+1}(0,\ell_t,b_t) + (1-\theta)V_{t+1}(m_t,\ell_t,b_t)\right] \Big\}.$$

By the usual envelope result, quasilinear preferences in the CM imply that CM value functions are linear in asset holdings:

$$\frac{\partial V_t(m,\ell,b)}{\partial m} = \phi_t, \quad \frac{\partial V_t(m,\ell,b)}{\partial \ell} = \phi_t(1+i_{\ell,t}), \quad \frac{\partial V_t(m,\ell,b)}{\partial b} = \phi_t(1+i_{b,t}).$$

Moving these one period forward and plugging them back into the above equation allows to rewrite households' problem as (10).

A.2. Confirming that entrepreneurs repay loans in the steady state

In the entrepreneurs' problem (12), we have implicitly taken it as given that entrepreneurs choose to work in the DM and repay their loans; we will now verify that this is indeed the case. Inserting the constraint into the objective function, we can reformulate problem (12) as

$$\max_{\{k_t, q_t^e\}} \quad \rho_t q_t^e - c(q_t^e, k_t) - \frac{1+\tilde{i}_{\ell,t+1}}{1+\iota_{t+1}} k_t.$$

Given that entrepreneurs have invested k_t in the CM, they are willing to work in the DM and repay their loan if and only if

$$\max_{q_t^e} \quad \rho_t q_t^e - c(q_t^e, k_t) - \frac{1+\tilde{i}_{\ell,t+1}}{1+\imath_{t+1}} k_t \geq 0, \tag{A.1}$$

i.e. if and only if the maximised return from production is enough to cover the cost of the loan and the disutility of working. If (A.1) is violated, entrepreneurs are better of not working at all in the DM and defaulting on their entire loan. To confirm that (A.1) is slack in equilibrium, recall first that since f is homogeneous of degree one, c is also homogeneous of degree one. Using $\kappa_t \equiv k_t/q_t^e$, we can thus rewrite entrepreneurs' objective function as

$$q_t^e \left[\rho_t - c\left(1, \kappa_t\right) - \frac{1+\tilde{i}_{\ell,t+1}}{1+\imath_{t+1}} \kappa_t \right], \tag{A.2}$$

and we can rewrite entrepreneurs' FOC for q_t^e, (14), as

$$\rho_t = c\left(1, \kappa_t\right) - \kappa_t c_k \left(1, \kappa_t\right). \tag{A.3}$$

Inserting the FOCs (13) and (A.3) into the objective function (A.2), we obtain that the optimised value of the objective function equals zero. Given that k_t is chosen optimally, we thus have

$$\max_{q_t^e} \left\{ \rho_t q_t^e - c(q_t^e, k_t) - \frac{1+\tilde{i}_{\ell,t+1}}{1+\imath_{t+1}} k_t \right\} = 0,$$

which confirms our initial conjecture that (A.1) is slack. More precisely, (A.1) is *just* slack, with the Lagrange multiplier being zero while the constraint holds at equality. This reflects the fact that entrepreneurs make zero profits in equilibrium, which in turn results from the CRS property of $f(k,h)$.

A.3. The optimality conditions of the bank's problem

The Lagrangian of the bank's problem writes:

$$\begin{aligned} \mathcal{L} = \theta u\left(\frac{d_t^I}{p_t}\right) + \beta(1-\theta)\phi_{t+1} d_{t+1}^P - \phi_t(m_t^b + a_t^b) \\ + \beta\phi_{t+1}\left[\zeta_t(m_t - \theta d_t^I) + \xi_t(m_t + a_t^b(1+i_{t+1}) - \theta d_t^I - (1-\theta) d_{t+1}^P) \right. \\ \left. + (1-\theta)\psi_t(d_{t+1}^P - d_t^I)\right]. \end{aligned}$$

The first-order conditions are

$$d_t^I : \quad \frac{u'(q_t)}{p_t} = \beta\phi_{t+1}\left(\zeta_t + \xi_t + \left(\tfrac{1}{\theta} - 1\right)\psi_t\right), \tag{A.4}$$

$$d_{t+1}^P : \quad \xi_t - \psi_t = 1, \tag{A.5}$$

$$m_t^b : \quad 1+\imath_t = \zeta_t + \xi_t, \tag{A.6}$$

$$a_t^b : \quad \xi_t = \frac{1+\imath_t}{1+i_{t+1}}, \tag{A.7}$$

where we have used $1+\imath_t \equiv \phi_t/(\beta\phi_{t+1})$. From (A.5)–(A.7), we obtain the expressions for the Lagrange multipliers in (21). Inserting the expressions for the multipliers into (A.4) yields Eq. (22). Finally, the bank's asset demand schedule (24) follows from the complementary slackness conditions $\zeta_t(m_t^b - \theta d_t^I) = 0$ and $\psi_t(d_{t+1}^P - d_t^I) = 0$.

A.4. Proof that a stationary monetary equilibrium exists

As shown in the main text, i pins down κ, which in turn pins down ρ. Furthermore, ρ and i together pin down q. To show that an SME in the banking economy always exists, it remains to show that there always exists an $i \in [0, \imath]$ such that condition (28) is fulfilled. Note that $\mathcal{A}, \rho$ and q can all be expressed as continuous functions of i. From (28), we have that an equilibrium with $i \in (0, \imath)$ exists iff:

$$Q(i) = 1 + i - \frac{(1-\theta)(1+\imath)\rho(i)q(i)}{\mathcal{A}(i)} = 0. \tag{A.8}$$

We also have from (28) that $Q(0) > 0$ and $Q(\imath) < 0$ if neither an equilibrium with $i = 0$ nor one with $i = \imath$ exists. From the intermediate value theorem, we then know that if there exists neither an equilibrium with $i = 0$ nor one with $i = \imath$, then an equilibrium with $i \in (0, \imath)$ exists, which proves that an equilibrium with $i \in [0, \imath]$ exists. As a side note, we have that $Q(0) < 0$ and $Q(\imath) > 0$ if both an equilibrium with $i = 0$ and one with $i = \imath$ exist. It then follows again from the intermediate value theorem that there also exists (at least one) equilibrium with $i \in (0, \imath)$.

A.5. Derivations of the thresholds for the penalty on early redemptions

A.5.1. Threshold to stop a run

Rewriting the condition in (42) gives

$$\eta d_S^I \leq (\eta + \sigma)\underline{d}_R^P(\eta, \sigma). \tag{A.9}$$

Since the right-hand side of (A.9) is strictly increasing in σ while the left-hand side does not change in σ, there is a unique threshold, denoted $\underline{\sigma}(\eta)$, such that condition (A.9) is fulfilled iff $\sigma \geq \underline{\sigma}(\eta)$. Substituting for $\underline{d}_R^P(\eta, \sigma)$ we can rewrite condition (A.9) as:

$$\begin{aligned} \eta d_S^I \leq (\eta+\sigma)\frac{(1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right)}{1-\lambda-\overline{\omega}(\eta)(\eta+\sigma)} \\ \Leftrightarrow \sigma\left[(1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right) + \overline{\omega}(\eta)\eta d_S^I\right] \geq (1-\lambda)\eta d_S^I \\ - \eta\left[(1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right) + \overline{\omega}(\eta)\eta d_S^I\right] \\ \Leftrightarrow \sigma \geq \left[\frac{d_S^I - m_S^b - (1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right)}{m_S^b - \lambda d_S^I + (1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right)}\right]\eta \equiv \underline{\sigma}(\eta). \end{aligned} \tag{A.10}$$

Notice that $\underline{\sigma}(0) = 0$. Also, $\underline{\sigma}(\eta)$ is strictly increasing in η iff $d_S^I \geq (1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right) + m_S^b$, which is the condition for banks to be fragile under pure deposit freezes (see (40)).

Furthermore, the fact that we must have $\sigma \leq 1 - \eta$ means there is an upper bound on η, denoted $\tilde{\eta}^{max}$, defined as the unique value of η solving $\underline{\sigma}(\eta) = 1 - \eta$. Note that setting $\eta \leq \tilde{\eta}^{max}$ is a necessary condition to stop a run. Solving for $\tilde{\eta}^{max}$ gives

$$\begin{aligned} \tilde{\eta}^{max} &= \frac{(1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right) + m_S^b - \lambda d_S^I}{(1-\lambda)d_S^I} \\ &= \overline{\eta} + \frac{(1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right)}{(1-\lambda)d_S^I} \in (\overline{\eta}, 1), \end{aligned} \tag{A.11}$$

where $\tilde{\eta}^{max} < 1$ follows from the fact that $d_S^I \geq (1+i_S)\left(b_S^b + \frac{m_S^b}{d_S^I}\ell_S^b\right) + m_S^b$. It follows that a run is stopped if (η, σ) satisfies $\eta \in [\overline{\eta}, \tilde{\eta}^{\max}]$ and $\sigma \in [\underline{\sigma}(\eta), 1-\eta]$.

A.5.2. Threshold to satisfy impatient depositors' incentive constraint

Note first that we can reformulate condition (45) as

$$u(\eta q_S) \geq \frac{\phi}{1+\imath}(\eta+\sigma)\overline{d}_R^P(\eta, \sigma). \tag{A.12}$$

Since the right-hand side of (A.12) is strictly increasing in σ while the left-hand side does not change in σ, there exists a unique threshold, denoted $\overline{\sigma}(\eta)$, such that condition (A.12) is fulfilled iff $\sigma \leq \overline{\sigma}(\eta)$. It will be useful to define:

$$T(\eta) \equiv \frac{m_S^b + (1+i_S)(b_S^b + (\lambda + (1-\lambda)\eta)\ell_S^b) - (\lambda + \theta(1-\lambda)\eta)d_S^I}{1-\lambda}. \tag{A.13}$$

In words, $T(\eta)$ denotes banks' per capita CM revenue in a run, given that the run is stopped and given that the impatient depositors who were not among the first λ to arrive withdraw fraction η of their deposit. 'Per capita' means that the total revenue is divided by the measure of depositors who did not manage to redeem before the bank imposed the redemption penalty. Notice that we have $T'(\eta) = (1+i_S)\ell^b_S - \theta d^I_S < 0$.[37] Substituting for $\overline{d}^P_R(\eta,\sigma)$, we can rewrite condition (A.12) as:

$$\begin{aligned} u(\eta q_S) &\geq \frac{(\eta+\sigma)}{(1-\theta(\eta+\sigma))}\frac{\phi}{1+\imath}T(\eta) \\ \Leftrightarrow \sigma\left[\theta u(\eta q_S) + \frac{\phi}{1+\imath}T(\eta)\right] &\leq (1-\theta\eta)u(\eta q_S) - \frac{\phi}{1+\imath}\eta T(\eta) \\ \Leftrightarrow \sigma &\leq \frac{u(\eta q_S)}{\theta u(\eta q_S) + \frac{\phi}{1+\imath}T(\eta)} - \eta \equiv \overline{\sigma}(\eta). \end{aligned} \tag{A.14}$$

Note that we have $\overline{\sigma}(0) = 0$ and

$$\overline{\sigma}'(\eta) = \frac{\frac{\phi}{1+\imath}\left(q_S T(\eta)u'(\eta q_S) - u(\eta q_S)T'(\eta)\right)}{\left(\theta u(\eta q_S) + \frac{\phi}{1+\imath}T(\eta)\right)^2} - 1, \tag{A.15}$$

with $\lim_{\eta\to 0}\overline{\sigma}'(\eta) = \frac{q_S u'(0)}{\frac{\phi}{1+\imath}T(0)} - 1 = +\infty$, which results from the fact that $u(q)$ satisfies the Inada conditions. Notice that $\overline{\sigma}'(\eta)$ need not be positive everywhere on $\eta \in (0,1)$. The reason is that the return to banks' asset portfolio increases in η due to lower loan defaults, which makes it more costly to give up a given share of the deposit. If this effect is strong enough, the redemption penalty which impatient depositors are willing to pay may decrease in η.

Next, we will show that $\overline{\sigma}''(\eta) < 0$, i.e. $\overline{\sigma}(\eta)$ is strictly concave on $\eta \in [0,1]$. We have $\overline{\sigma}''(\eta) = \frac{f'g - fg'}{g^2}$, where f and g are defined by $\overline{\sigma}'(\eta) \equiv \frac{f}{g} - 1$, and f' and g' denote the derivatives of f and g w.r.t. η. We have

$$f' = \underbrace{\frac{\phi}{1+\imath}q_S^2 T(\eta)}_{>0}\underbrace{u''(\eta q_S)}_{<0} < 0 \quad \text{and}$$

$$g' = 2\underbrace{\left[\theta u(\eta q_S) + \frac{\phi}{1+\imath}T(\eta)\right]}_{>0}\underbrace{\left[\theta q_S u'(\eta q_S) + \frac{\phi}{1+\imath}T'(\eta)\right]}_{>0 \text{ for } \eta\in[0,1]} > 0,$$

from which it follows that $\overline{\sigma}''(\eta) < 0$. To see why $g' > 0$, note that

$$\begin{aligned} &\theta q_S u'(\eta q_S) + \frac{\phi}{1+\imath}T'(\eta) > 0 \\ &\Leftrightarrow \theta q_S u'(\eta q_S) > \frac{\phi}{1+\imath}(\theta d^I_S - (1+i_S)\ell^b_S) \\ &\Leftrightarrow \frac{u'(\eta q_S)}{\rho_S} > 1 - \frac{\phi}{1+\imath}\frac{(1+i_S)\ell^b_S}{\theta\rho_S q_S}, \end{aligned} \tag{A.16}$$

where we used $T'(\eta) = -(\theta d^I_S - (1+i_S)\ell^b_S)$, $d^I_S = p_S q_S$ and $\rho_S \equiv \frac{\phi}{1+\imath}p_S$. Since $\frac{u'(q_S)}{\rho_S} > 1$ in the steady state (see (14) and (25)), we know that condition (A.16) is fulfilled for any $\eta \in [0,1]$.

[37] To see why the derivative is negative, recall that θd^I_S equals the total steady state DM revenue of entrepreneurs while $(1+i_S)\ell^b$ equals their total loan repayment in steady state. Since entrepreneurs keep part of their revenue as compensation for their labour effort, we have $\theta d^I_S > (1+i_S)\ell^b_S$.

A.5.3. Threshold to prevent a run

Substituting for $\overline{d}^P_R(\eta,\sigma)$ we can rewrite condition (46) as:

$$\begin{aligned} (1-\lambda)(1-\theta(\eta+\sigma))d^I_S &\leq m^b_S + (1+i_S)(b^b_S + (\lambda+(1-\lambda)\eta)\ell^b_S) \\ &\quad - ((\lambda+\theta(1-\lambda)\eta)d^I_S) \\ \Leftrightarrow \sigma\theta(1-\lambda)d^I_S &\geq (\lambda+(1-\lambda))d^I_S - \left[m^b_S + (1+i_S)b^b_S\right. \\ &\quad \left. + (1+i_S)(\lambda+(1-\lambda)\eta)\ell^b_S\right] \\ \Leftrightarrow \sigma &\geq \left[\frac{d^I_S - (m^b_S + (1+i_S)(b^b_S + \lambda\ell^b_S))}{(1-\lambda)\theta d^I_S}\right] \\ &\quad - \frac{(1+i_S)\ell^b_S}{\theta d^I_S}\eta \equiv \underline{\hat{\sigma}}(\eta). \end{aligned} \tag{A.17}$$

Note that we have $\underline{\hat{\sigma}}(0) \in (0,1)$ and $\underline{\hat{\sigma}}'(\eta) = -\frac{(1+i_S)\ell^b_S}{\theta d^I_S} < 0$, with $|\underline{\hat{\sigma}}'(\eta)| < 1$.[38]

A.6. Proof of Proposition 8

Given that banks set (η,σ) such as to stop the run, DM output in a run is maximised by setting η as high as possible subject to the relevant constraints. Since we start from the premise that runs cannot be prevented, we can disregard constraint (46). The relevant constraints – besides banks' liquidity constraint – are thus (42) and (45). We are therefore looking for values (η,σ) that maximise η subject to the constraint that $\underline{\sigma}(\eta) \leq \sigma \leq \overline{\sigma}(\eta)$ with $\sigma \in [0, 1-\eta]$, and subject to the liquidity constraint $\eta \leq \overline{\overline{\eta}}$.

Recall from Appendices A.5.1 and A.5.2 that $\underline{\sigma}(0) = \overline{\sigma}(0) = 0$, where $\underline{\sigma}(\eta)$ is linearly increasing in η, while $\overline{\sigma}(\eta)$ is strictly concave on $\eta \in [0,1]$ with $\lim_{\eta\to 0}\overline{\sigma}'(\eta) = +\infty$. This implies that whenever there exist values $\eta \in [0,1]$ for which $\overline{\sigma}(\eta) < \underline{\sigma}(\eta)$, then there exists a unique strictly positive value $\hat{\eta}^{max}$ such that $\overline{\sigma}(\eta) \geq \underline{\sigma}(\eta)$ iff $\eta \leq \hat{\eta}^{max}$. If $\overline{\sigma}(\eta) \geq \underline{\sigma}(\eta)$ for all $\eta \in [0,1]$, we define $\hat{\eta}^{max} = 1$. Next, we will show that $\hat{\eta}^{max} > \overline{\eta}$. We have

$$\begin{aligned} &\overline{\sigma}(\overline{\eta}) > \underline{\sigma}(\overline{\eta}) \\ \Leftrightarrow\quad &\frac{1}{\theta + \frac{\phi}{1+\imath}\frac{T(\overline{\eta})}{u(\overline{\eta}q_S)}} > \frac{\overline{\eta}(1-\lambda)d^I_S}{(1+i_S)\left(b^b_S + \frac{m^b_S}{d^I_S}\ell^b_S\right) + m^b_S - \lambda d^I_S} \\ \Leftrightarrow\quad &T(\overline{\eta}) > \frac{\phi}{1+\imath}\frac{\overline{\eta}d^I_S T(\overline{\eta})}{u(\overline{\eta}q_S)} \\ &\frac{u(\overline{\eta}q_S)}{\overline{\eta}q_S} > \rho_S, \end{aligned} \tag{A.18}$$

where we used $d^I_S = p_S q_S \equiv \frac{1+\imath}{\phi}\rho_S q_S$ in the last step. Since $u(q)$ is strictly concave and $u(0) = 0$, we have $u(q)/q > u'(q)$. Since $u'(q_S) > \rho_S$ (see (14) and (25)), condition (A.18) is fulfilled, from which it follows that $\hat{\eta}^{max} > \overline{\eta}$.

Finally, from Appendix A.5.1, we know that the highest value of η consistent with $\underline{\sigma}(\eta) \leq 1-\eta$, i.e. consistent with stopping a run, equals $\tilde{\eta}^{\max}$, with $\tilde{\eta}^{\max} \in (\overline{\eta}, 1)$. It then follows that constraints (42) and (45) can only be jointly satisfied if $\eta \leq \min\{\tilde{\eta}^{max}, \hat{\eta}^{max}\} \in (\overline{\eta}, 1)$. The fact that banks' liquidity constraint requires additionally that $\eta \leq \overline{\overline{\eta}}$ then leads to the result in Proposition 8.

A.7. Example with $\lambda = 0$ where runs cannot be prevented

Fig. 2 shows an example where there exists no (η,σ) that prevents runs, despite the fact that banks can react to runs immediately ($\lambda = 0$). The functional forms used here are the same as in the example shown in

[38] See Footnote for why the absolute value of the derivative is less than one.

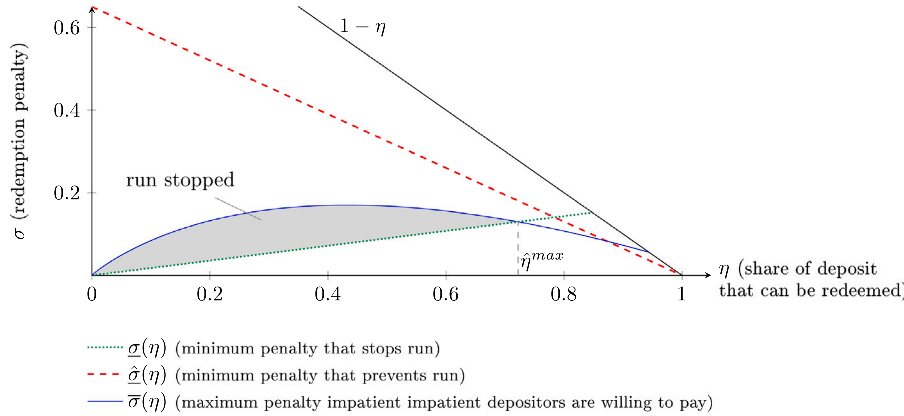

Fig. 2. Example where runs cannot be prevented.

Table 2
Parameter values for Fig. 2.

α	ν	θ	n	λ	ι	$\mathcal{B}$	q_S	ρ_S	κ_S
0.65	0.95	0.6	0.6	0	0.02	0	$5.725 * 10^{-7}$	1.886	1.251

Fig. 1, and the parameter values are given in Table 2. As in the example of Fig. 1, parameters imply that the economy is in a ZLB equilibrium.

We can see graphically that the set of values (η, σ) satisfying constraints (42), (45) and (46) is empty. Notice that the banks' liquidity constraint is fulfilled for any $\eta \in [0,1]$ since banks can react to runs without delay. What is key is that impatient depositors' willingness to pay a redemption penalty is low, which results from the fact that the DM utility function is close to linear. In particular, the redemption penalty cannot be set high enough to deter patient depositors from running, since impatient depositors would not be willing to incur such a penalty. Nevertheless, banks can stop runs by partially freezing deposits and charging a modest penalty on redemptions once a run has started (grey area). DM output in a run is maximised by setting $(\eta, \sigma) = (\hat{\eta}^{max}, \underline{\sigma}(\hat{\eta}^{max}))$, where $\hat{\eta}^{max}$ corresponds to the highest DM payout consistent with stopping the run.

A.8. Proof of Lemma 3

From DM market clearing (30), we have that

$$\frac{1+\pi_S}{1+\pi}[\theta+\eta(1-\theta)] = \min\left\{\frac{\alpha(\rho)}{\alpha(\rho_S)}[\theta+\eta(1-\theta)], 1\right\}\frac{\rho q^e(\rho)}{\rho_S q^e(\rho_S)}, \tag{A.19}$$

where we used (31) to substitute for χ and (52) to substitute for M^{RE}. Recall that η is fixed given the discounts (δ_b, δ_ℓ) (see (50)). Next, as seen in Section 5, $q^e(\rho)$ is continuous and strictly increasing in ρ. Furthermore, from (33), we have $\rho\alpha(\rho)q^e(\rho) = \max_{q^e}\left\{\rho q^e - c(q^e, k_S)\right\}$, and hence, by the envelope condition, $\partial[\rho\alpha(\rho)q^e(\rho)]/\partial\rho = q^e(\rho) > 0$. It follows from this that the RHS of (A.19) is continuous and strictly increasing in ρ (with a kink at $\alpha(\rho)[\theta+\eta(1-\theta)]/\alpha(\rho_S) = 1$), which in turn means that (A.19) implicitly defines ρ as a strictly decreasing function of π.

A.9. Proof of Proposition 9

We start by defining and characterising the threshold $\bar{\delta}_\ell$. To do so, consider first the equality

$$\delta_\ell = \zeta\left[\theta+(1-\theta)\eta(\delta_\ell,\delta_b)\right] = \zeta \min\left\{1, \frac{m_S^b + (1+i_S)(\delta_\ell \ell_S^b + \delta_b b_S^b)}{d_S^I}\right\}, \tag{A.20}$$

where we used (50) to substitute for η and made explicit the dependence of η on (δ_ℓ, δ_b). Denote

$$\hat{\delta}_\ell(\zeta) \equiv \frac{\zeta[m_S^b + (1+i_S)\delta_b b_S^b]}{d_S^I - \zeta(1+i_S)\ell_S^b} \tag{A.21}$$

as the unique, strictly positive value of δ_ℓ solving $\delta_\ell = \zeta[m_S^b + (1+i_S)(\delta_\ell \ell_S^b + \delta_b b_S^b)]/d_S^I$. Note that both the numerator and the denominator in (A.21) are strictly positive since (i) $m_S^b > 0$ and (ii) total loan repayments in the steady state, $(1+i_S)\ell_S^b$, cannot exceed total DM spending in the steady state, θd_S^I. We then define $\bar{\delta}_\ell$ as the unique value of δ_ℓ solving (A.20), which equals:

$$\bar{\delta}_\ell(\zeta) \equiv \begin{cases} \hat{\delta}_\ell(\zeta) & \text{if } \hat{\delta}_\ell(\zeta) \le \zeta \\ \zeta & \text{if } \hat{\delta}_\ell(\zeta) > \zeta. \end{cases} \tag{A.22}$$

Notice that $\bar{\delta}_\ell(\zeta)$ is strictly increasing in ζ, and that we have $\delta_\ell < \bar{\delta}_\ell \Leftrightarrow \delta_\ell < \zeta[\theta+\eta(1-\theta)]$ as well as $\delta_\ell > \bar{\delta}_\ell \Leftrightarrow \delta_\ell > \zeta[\theta+\eta(1-\theta)]$.

To show that $\bar{\delta}_\ell(1) = 1$, note first that we have from (A.21) that

$$\hat{\delta}_\ell(1) = \frac{m_S^b + (1+i_S)\delta_b b_S^b}{d_S^I - (1+i_S)\ell_S^b} \ge \frac{m_S^b + (1+i_S)\,\hat{\delta}_\ell(1)\,b_S^b}{d_S^I - (1+i_S)\ell_S^b}, \tag{A.23}$$

where the second inequality results from the fact that $\delta_b \ge \delta_\ell$. Solving (A.23) for $\hat{\delta}_\ell(1)$ yields

$$\hat{\delta}_\ell(1) \ge \frac{m_S^b}{d_S^I - (1+i_S)(\ell_S^b + b_S^b)} \ge 1, \tag{A.24}$$

where the last inequality is due to the fact that $d_S^I \le m_S^b + (1+i_S)(b_S^b + \ell_S^b)$, which in turn follows from patient depositors' IC constraint, $d_S^I \le d_S^P$. From (A.22) and (A.24), it follows that $\bar{\delta}_\ell(1) = 1$.

The next step is to show that $\delta_\ell \le \bar{\delta}_\ell(\zeta)$ rules out a decrease in the capital share, $\alpha(\rho)$, relative to the steady state:

Show that $\delta_\ell \le \bar{\delta}_\ell(\zeta) \Rightarrow \alpha(\rho) \ge \alpha(\rho_S)$. We proceed with a proof by contradiction and suppose that $\delta_\ell \le \bar{\delta}_\ell(\zeta)$ but $\alpha(\rho) < \alpha(\rho_S)$. Note first that entrepreneurs' real revenue, $\rho q^e(\rho)$, is strictly decreasing in π since ρ is strictly decreasing in π (see Lemma 3) while $q^e(\rho)$ is strictly increasing in ρ (see Section 5). It then follows from (34) that

$$\frac{\alpha(\rho_S)}{\alpha(\rho)} \le \frac{1+\pi}{1+\pi_S} \tag{A.25}$$

for any $\pi \ge \pi_S$. Next, we use (31), (52) and (58) to obtain:

$$\frac{1+\pi}{1+\pi_S} = 1 + \max \times \left\{\frac{(1+i_S)\left(\delta_\ell - \zeta\min\left\{\frac{\alpha(\rho)}{\alpha(\rho_S)}[\theta+\eta(1-\theta)], 1\right\}\right)\tau_\ell \ell_S^b - (1+i_S)(1-\delta_b)\tau_b b_S^b}{M_S + (1+i_S)B_S}, 0\right\}. \tag{A.26}$$

Taking the derivative of the RHS in (A.26) with respect to $\alpha(\rho_S)/\alpha(\rho)$ yields

$$\frac{\partial[(1+\pi)/(1+\pi_S)]}{\partial[\alpha(\rho_S)/\alpha(\rho)]} \in \left\{ \frac{\zeta(1+i_S)\tau_\ell \ell_S^b}{M_S+(1+i_S)B_S}\left(\frac{\alpha(\rho)}{\alpha(\rho_S)}\right)^2(\theta+\eta(1-\theta)), 0\right\}. \tag{A.27}$$

Note that we have $(1+i_S)\ell_S^b < M_S$, i.e. aggregate nominal loan repayments in the steady state are strictly lower than the steady state money stock since entrepreneurs keep part of their revenue as compensation for their labour effort. Since we have $\zeta \le 1, \tau_\ell \le 1$ and $\eta \le 1$, we obtain from (A.27) that

$$\frac{\partial[(1+\pi)/(1+\pi_S)]}{\partial[\alpha(\rho_S)/\alpha(\rho)]} < 1 \quad \text{for} \quad \frac{\alpha(\rho_S)}{\alpha(\rho)} \ge 1. \tag{A.28}$$

Furthermore, for any $\delta_\ell \le \bar{\delta}_\ell(\zeta)$, which as seen above is equivalent to $\delta_\ell \le \zeta[\theta+\eta(1-\theta)]$, we obtain from (A.26) that $\pi = \pi_S$ if $\alpha(\rho) = \alpha(\rho_S)$. This means that if $\delta_\ell \le \bar{\delta}_\ell(\zeta)$, then (A.25) holds at equality if $\alpha(\rho_S)/\alpha(\rho) = 1$. It then follows immediately from (A.28) that (A.25) is violated for any $\alpha(\rho_S)/\alpha(\rho) > 1$. Thus, assuming $\delta_\ell \le \bar{\delta}_\ell(\zeta)$ and $\alpha(\rho_S)/\alpha(\rho) > 1$ leads to a contradiction.

Intuitively, condition (A.25) states that to maintain entrepreneurs' incentives to produce in the DM, a decrease in the capital share (the share of entrepreneurs' revenue left after compensating them for their labour effort) must go in hand with a sufficiently strong increase in inflation reducing entrepreneurs' real debt burden. At the same time, (A.28) says that a fall in the capital share by itself does not lead to sufficiently large losses for the government to create such an increase in inflation. Put differently, a fall in the capital share can only occur in equilibrium if the government's losses on its loan purchases are such that inflation increases even without a fall in the capital share.

To show that $\delta_\ell > \zeta\chi$ if and only if $\delta_\ell > \bar{\delta}_\ell(\zeta)$, we first prove the 'if' statement and then the 'only if' statement:

Show that $\delta_\ell > \bar{\delta}_\ell(\zeta) \Rightarrow \delta_\ell > \zeta\chi$. We proceed with a proof by contradiction and suppose that $\delta_\ell > \bar{\delta}_\ell$ but $\delta_\ell \le \zeta\chi$. From (58), we have that $\delta_\ell \le \zeta\chi$ implies $\pi = \pi_S$ and hence, by Lemma 3, $\rho = \rho_S$. From (31) and (52), we then have $\chi = \theta + \eta(1-\theta)$. Therefore, $\delta_\ell \le \zeta\chi$ implies $\delta_\ell \le \zeta[\theta+\eta(1-\theta)]$, which, as seen further above, means $\delta_\ell \le \bar{\delta}_\ell$ and thus leads to a contradiction.

Show that $\delta_\ell \le \bar{\delta}_\ell(\zeta) \Rightarrow \delta_\ell \le \zeta\chi$. We again proceed with a proof by contradiction and suppose that $\delta_\ell \le \bar{\delta}_\ell$ but $\delta_\ell > \zeta\chi$. Note first that if $\delta_\ell \le \bar{\delta}_\ell$ and $\delta_\ell > \zeta\chi$, then we have $\zeta\chi < \delta_\ell \le \bar{\delta}_\ell \le \zeta$, which implies $\chi < 1$. Using (31) to substitute for χ in $\delta_\ell > \zeta\chi$ then yields

$$\frac{\alpha(\rho_S)}{\alpha(\rho)} > \frac{\zeta}{\delta_\ell}\frac{M^{RE}}{\theta d_S^I} = \frac{\zeta[\theta+\eta(1-\theta)]}{\delta_\ell}, \tag{A.29}$$

where we used (52) to substitute for M^{RE}. Since $\delta_\ell \le \bar{\delta}_\ell$ implies $\delta_\ell \le \zeta[\theta+\eta(1-\theta)]$, we have from (A.29) that $\alpha(\rho) < \alpha(\rho_S)$, which leads to a contradiction because, as seen above, $\alpha(\rho) < \alpha(\rho_S)$ cannot occur when $\delta_\ell \le \bar{\delta}_\ell(\zeta)$.

Finally, given that $\delta_\ell > \zeta\chi \Leftrightarrow \delta_\ell > \bar{\delta}_\ell(\zeta)$, it follows immediately from (58) that $\pi = \pi_S$ if $\delta_\ell \le \bar{\delta}_\ell(\zeta)$.

A.10. Proof of Proposition 11

The proof of the first part of Proposition 11 about ZLB and PLI equilibria is contained in the main text. Consider now FLI equilibria. Suppose for the moment that $\eta < 1$, in which case we have from (50) and (59) that

$$\chi = \frac{m_S^b + (1+i_S)(\delta_b b_S^b + \delta_\ell \ell_S^b)}{d_S^I}.$$

Inserting this into condition (56), and using the fact that $m_S^b = \theta d_S^I$ and $i_S = \imath$ in FLI equilibria, yields

$$(1-\theta)d_S^I \le (1+\imath)b_S^b + (1+\imath)\left(\theta + \frac{(1+\imath)(\delta_b b_S^b + \delta_\ell \ell_S^b)}{d_S^I}\right)\ell_S^b. \tag{A.30}$$

Substituting $\phi d_S^I = \phi\rho_S q_S \equiv (1+\imath)\rho_S q_S$, $\phi b_S^b = \mathcal{B}$ and $\phi \ell_S^b = \theta q_S \kappa_S$ in condition (A.30) (the latter two equations follow from the fact that banks hold all assets in the economy if $\mathcal{A}_S \le \overline{\mathcal{A}}$) gives

$$(1-\theta)\rho_S q_S \le \mathcal{B} + \theta^2 q_S\kappa_S + \frac{\theta\kappa_S}{\rho_S}(\delta_b\mathcal{B} + \delta_\ell\theta q_S\kappa_S). \tag{A.31}$$

Using $\mathcal{A}_S = \mathcal{B} + \theta q_S\kappa_S$ and the fact that $\kappa_S = \kappa^*$ in FLI equilibria, condition (A.31) can then further be rewritten as

$$(\delta_\ell - 1)\theta q_S\kappa^* + (\delta_b - 1)\mathcal{B} \ge -\left(1 + \frac{\rho_S}{\theta\kappa}^*\right)(\mathcal{A}_S - (1-\theta)\rho_S q_S). \tag{A.32}$$

So far we have taken it as given that $\eta < 1$, i.e. the discounts are high enough such that banks impose a redemption penalty in the DM. Inserting $m_S^b = \theta d_S^I$ into condition (51) and substituting for d_S^I, ℓ_S^b and b_S^b in the same manner as above yields that in an FLI equilibrium

$$\eta < 1 \;\Leftrightarrow\; (\delta_\ell - 1)\theta q_S\kappa^* + (\delta_b - 1)\mathcal{B} < -[\mathcal{A}_S - (1-\theta)\rho_S q_S].$$

We can see immediately that $\eta < 1$ whenever condition (A.32) is violated, which means that (A.32) is both necessary and sufficient for emergency liquidity to prevent runs.

Finally, inserting $\delta_b = 1$ and $\delta_\ell = \bar{\delta}_\ell(\zeta)$ into condition (A.32) and rearranging terms yields

$$\bar{\delta}_\ell(\zeta) \ge 1 - \frac{\theta\kappa^* + \rho_S}{(\theta\kappa^*)^2 q_S}\left[\mathcal{A}_S - (1-\theta)\rho_S q_S\right]. \tag{A.33}$$

The result in (60) then follows from (A.33) together with the fact that $\bar{\delta}_\ell(\zeta)$ is strictly increasing in ζ.

A.11. Proof of Proposition 12

We first show that if $\delta_\ell = \delta_b = 1$ and $\alpha'(\rho) \le 0$, then $\chi = 1$. We proceed with a proof by contradiction and suppose that $\delta_\ell = \delta_b = 1$, $\alpha'(\rho) \le 0$ and $\chi < 1$. From (31), we then have that

$$\chi = \frac{\alpha(\rho)}{\alpha(\rho_S)}\frac{M^{RE}}{\theta d_S^I} = \frac{\alpha(\rho)}{\alpha(\rho_S)}[\theta+\eta(1-\theta)] = \frac{\alpha(\rho)}{\alpha(\rho_S)}, \tag{A.34}$$

where we used (52) to substitute for M^{RE}, and we used the fact that $\eta = 1$ when $\delta_\ell = \delta_b = 1$ (see (50)). Since ρ is strictly decreasing in π (see Lemma 3), we have $\alpha(\rho) \ge \alpha(\rho_S)$ for any $\pi \ge \pi_S$ if $\alpha(\rho)$ is weakly decreasing in ρ. Therefore, we have from (A.34) that $\chi \ge 1$, which leads to a contradiction. Since we know from Lemma 2 that runs are prevented if $\chi = 1$, this proves the first part of Proposition 12.

We now prove the second part of Proposition 12, which states that runs cannot be prevented if $\alpha'(\rho) > 0$, $\zeta < 1$ and the economy is in a ZLB or PLI equilibrium. From Lemma 2, we know that runs are only prevented in ZLB and PLI equilibria if defaults in a run are avoided completely ($\chi = 1$). We show with a proof by contradiction that this is not possible. Suppose $\alpha'(\rho) > 0$, $\zeta < 1$, the economy is a ZLB or PLI equilibrium, and $\chi = 1$. From (31), we have that $\chi = 1$ requires

$$\frac{\alpha(\rho)}{\alpha(\rho_S)}\frac{M^{RE}}{\theta d_S^I} \ge 1 \quad\Leftrightarrow\quad \frac{\alpha(\rho)}{\alpha(\rho_S)}[\theta+\eta(1-\theta)] \ge 1, \tag{A.35}$$

where we used (52) to substitute for M^{RE}. Since $\eta \le 1$, and ρ is strictly decreasing in π (see Lemma 3), condition (A.35) can only be fulfilled if $\pi = \pi_S$ and $\eta = 1$, given that $\alpha'(\rho) > 0$. We know from Lemma 1 that $\eta = 1$ requires $\delta_\ell = \delta_b = 1$ in ZLB and PLI equilibria, and we know from Proposition 9 that $\zeta\chi < 1$ if $\zeta < 1$. It then follows from (58) that if $\delta_\ell = \delta_b = 1$, we have $\pi > \pi_S$, such that we arrive at a contradiction.

A.12. Proof of Proposition 13

With real loan contracts, entrepreneurs' real debt burden is fixed at its steady state level, which, as shown in the proof of Proposition 2, means that we must have $\rho = \rho_S$. Suppose now the government sets $\delta_\ell = \delta_b = 1$, in which case we have $\eta = 1$ (see (50)). Given $\rho = \rho_S$, we then have from DM market clearing (30) that

$$\chi = \frac{1+\pi_S}{1+\pi}\frac{M^{RE}}{\theta d_S^I} = \frac{1+\pi_S}{1+\pi}[\theta+\eta(1-\theta)] = \frac{1+\pi_S}{1+\pi}, \tag{A.36}$$

where we have used (52) to substitute for M^{RE}. Eq. (A.36) states that an increase in inflation – with the associated fall in impatient depositors' real purchasing power – will lead to some defaults when loan contracts are real. Next, we have that

$$\chi \frac{1+\pi}{1+\pi_S}(1+i_S)\ell^b_S = (1+i_S)\ell^b_S, \quad (A.37)$$

where the LHS equals the nominal return on banks' loan portfolio with real loan contracts, and we have used (A.36) to substitute for χ. Eq. (A.37) shows that the nominal return which banks earn on their loan portfolio in case of a run is the same as in the steady state. This means that banks can always pay the promised amount to those withdrawing in the CM, such that patient depositors have no incentive to run on the banks in the first place.

A.13. Reinterpreting the government's intervention as secured lending

Suppose that, instead of purchasing assets outright, the government stands ready to provide emergency credit at a gross interest rate of $1/\delta_b$ when the loan is secured by government bonds and at gross interest rate of $1/\delta_\ell$ when it is secured by loans, where we continue assuming $\delta_\ell \leq \delta_b \leq 1$. Banks need to post collateral with a gross face value (i.e. principal plus interest) equal to the gross face value of the emergency loan. That is, to obtain one unit of money in the DM, banks need to post either bonds with a gross face value of $1/\delta_b$ or loans with a gross face value of $1/\delta_\ell$ (or any combination thereof). Analogous to Section 7, τ_b and τ_ℓ denote the fraction of bonds and loans, respectively, that banks pledge as collateral. Emergency loans extended to banks are due in the next CM. If banks fail to repay the due amount in full, the government seizes the collateral.

As before, we assume an inefficiency on the side of the government when it comes to emergency loans that are collateralised by loans. Specifically, we reinterpret ζ as reflecting a cost incurred by the government when managing the loans pledged as collateral, with the cost being proportional to the realised value of the pledged loans. That is, if banks post an amount ℓ of loans as collateral, the government incurs a cost of $(1-\zeta)\chi(1+i_S)\ell$. We also continue assuming that the government cannot increase its real indebtedness, i.e. constraint (48) applies. In said constraint, $M + (1+i_S)B$ now denotes the government's nominal liabilities at the beginning of the next CM, just after entrepreneurs have repaid loans and banks have repaid their emergency credit. The provision of emergency liquidity will then be inflationary whenever it leads to an increase in nominal government liabilities beyond the point at which the assets pledged as collateral mature.

It is easy to see that the total amount of liquidity a bank can access by pledging all its assets as collateral is still given by (49). This implies that the redemption penalty banks need to charge in order to stop a run is still given by (50), and the aggregate cash held by impatient depositors in a run equals (52). Eq. (53) remains the same as well, with the RHS now denoting excess reserves plus the total money a bank raises by obtaining a secured credit from the government. As before, we assume $\tau_b = \tau_\ell = \theta + (1-\theta)\mathcal{I}$, i.e. banks pledge proportional amounts of bonds and loans when obtaining an emergency credit. To consider the effect of emergency liquidity on net nominal government liabilities, note first that the gross face value of the emergency loan equals the gross face value of the pledged collateral,

$$\tau_b(1+i_S)b^b_S + \tau_\ell(1+i_S)\ell^b_S,$$

while the amount lent by the government in the DM is

$$\tau_b\delta_b(1+i_S)b^b_S + \tau_\ell\delta_\ell(1+i_S)\ell^b_S.$$

Banks will default on their emergency loans whenever $\chi < 1$ as the value of the pledged collateral is then lower than the amount due on the emergency loans. The government's CM revenue from its emergency loan is therefore

$$\tau_b(1+i_S)b^b_S + \chi\tau_\ell(1+i_S)\ell^b_S.$$

The government's net profit or loss from the provision of emergency loans equals

$$\Pi = \tau_b(1+i_S)(1-\delta_b)b^b_S + \tau_\ell(1+i_S)(\chi-\delta_\ell)\ell^b_S - \tau_\ell(1-\zeta)\chi(1+i_S)\ell^b_S,$$

where the last term is the cost incurred by managing the pledged loans. The government's nominal liabilities after the provision of emergency liquidity equal $M_S + (1+i_S)B_S - \Pi$, which is the same as (57). It follows that inflation created by the government's intervention is still given by (58), which means that all further results derived in Section 7 remain the same. Note in particular that a bank's CM payouts – and hence the incentives to run for patient depositors – do not depend on whether the bank sells a given amount of assets in the DM or pledges them as collateral.

References

Altermatt, L., van Buggenum, H., Voellmy, L., 2024. Money creation in a neoclassical economy: Equilibrium multiplicity and the liquidity trap. KOF Working Paper No. 516.
Andolfatto, D., Berentsen, A., Martin, F., 2020. Money, banking and financial markets. Rev. Econ. Stud. 87 (5), 2049–2086.
Andolfatto, D., Nosal, E., 2020. Shadow bank runs. FRB Atlanta Working Paper No. 2020-14.
Andrew, A.P., 1908. Hoarding in the panic of 1907. Q. J. Econ. 22 (2), 290–299.
Aruoba, S.B., Waller, C., Wright, R., 2011. Money and capital. J. Monetary Econ. 58 (2), 98–116.
Bagehot, W., 1873. Lombard Street. Henry S. King & Co.
Bech, M., Faruqui, U., Ougaard, F., Picillo, C., 2018. Payments are a-changin' but cash still rules. BIS Q. Rev. 67–80.
Berentsen, A., Camera, G., Waller, C., 2007. Money, credit and banking. J. Econom. Theory 135 (1), 171–195.
Carapella, F., 2015. Banking panics and deflation in dynamic general equilibrium. Finance and Economics Discussion Series 2015-018. Board of Governors of the Federal Reserve System.
Comptroller of the Currency, 1907. Annual Report of the Comptroller of the Currency. Washington Government Printing Office.
Daseking, C., Ghosh, A., Lane, T., Thomas, A., 2004. Lessons from the crisis in Argentina. IMF Occasional Paper 236.
Diamond, D.W., Dybvig, P.H., 1983. Bank runs, deposit insurance, and liquidity. J. Polit. Econ. 91 (3), 401–419.
Engineer, M., 1989. Bank runs and the suspension of deposit convertibility. J. Monetary Econ. 24 (3), 443–454.
Ennis, H., Keister, T., 2009. Bank runs and institutions: The perils of intervention. Amer. Econ. Rev. 99 (4), 1588–1607.
Ennis, H.M., Keister, T., 2010. Banking panics and policy responses. J. Monetary Econ. 57 (4), 404–419.
European Commission, 2013. The economic adjustment programme for cyprus. European Economy Occasional Papers 149.
Farhi, E., Tirole, J., 2020. Shadow banking and the four pillars of traditional financial intermediation. Rev. Econ. Stud. 88 (6), 2622–2653.
Friedman, M., Schwartz, A.J., 1963. A Monetary History of the United States. Princeton University Press.
Geromichalos, A., Herrenbrueck, L., 2022. The liquidity-augmented model of macroeconomic aggregates: A new monetarist DSGE approach. Rev. Econ. Dyn. 45, 134–167.
Gertler, M., Kiyotaki, N., 2015. Banking, liquidity, and bank runs in an infinite horizon economy. Amer. Econ. Rev. 105 (7), 2011–2043.
Goldstein, I., Kopytov, A., Shen, L., Xiang, H., 2020. Bank heterogeneity and financial stability. NBER Working Paper 27376.
Gorton, G., 1988. Banking panics and business cycles. Oxford Econ. Pap. 40 (4), 751–781.
Gorton, G., Huang, L., 2006. Bank panics and the endogeneity of central banking. J. Monetary Econ. 53 (7), 1613–1629.
Gu, C., Monnet, C., Nosal, E., Wright, R., 2023. Diamond-Dybvig and beyond: On the instability of banking. Eur. Econ. Rev. 154.
Gutierrez, J.G., Montes-Negret, F., 2004. Argentina's banking system: Restoring financial viability. World Bank Working Paper N.2/04.
James, J.A., McAndrews, J., Weiman, D.F., 2013. Wall Street and Main Street: the macroeconomic consequences of New York bank suspensions, 1866–1914. Cliometrica 7, 99–130.
Jaremski, M.S., Wheelock, D.C., 2023. Interbank networks and the interregional transmission of financial crises: Evidence from the panic of 1907. NBER Working Paper 31270.
Kashyap, A.K., Tsomocos, D.P., Vardoulakis, A.P., 2024. Optimal bank regulation in the presence of credit and run-risk. J. Polit. Econ. 132 (3), 772–823.
Keister, T., Mitkov, Y., 2023. Allocating losses: Bail-ins, bailouts and bank regulation. J. Econom. Theory 210 (C), 1–29.

Kiguel, M., 2011. Argentina's 2001 economic and financial crisis: Lessons for Europe. In: Think Tank 20: Beyond Macroeconomic Policy Coordination Discussions in the G-20. Brookings Institution, (Chapter 2).
Kiyotaki, N., Wright, R., 1989. On money as a medium of exchange. J. Polit. Econ. 97 (4), 927–954.
Lagos, R., Wright, R., 2005. A unified framework for monetary theory and policy analysis. J. Polit. Econ. 113 (3), 463–484.
Liu, X., 2019. Diversification and systemic bank runs. mimeo.
Martin, A., 2006. Liquidity provision vs. deposit insurance: Preventing bank panics without moral hazard. Econom. Theory 28 (1), 197–211.
Matta, R., Perotti, E.C., 2024. Pay, stay, or delay? How to settle a run. Rev. Financ. Stud. 37 (4), 1368–1407.
Peck, J., Setayesh, A., 2023. Bank runs and the optimality of limited banking. Rev. Econ. Dyn. 47, 100–110.
Reinhart, C.M., Rogoff, K.S., 2008. Is the 2007 US sub-prime financial crisis so different? An international historical comparison. Am. Econ. Rev.: Pap. Proc. 98 (2), 339–344.
Robatto, R., 2019. Systemic bank panics, liquidity risk, and monetary policy. Rev. Econ. Dyn. 34, 20–42.
Rochet, J., Vives, X., 2004. Coordination failures and the lender of last resort: Was bagehot right after all? J. Eur. Econom. Assoc. 2 (6), 1116–1147.
Sanches, D., 2018. Banking panics and output dynamics. Rev. Econ. Dyn. 29, 148–171.
Saxton, J., 2003. Argentina's economic crisis: Causes and cures. Joint Economic Committee United States Congress.
Schilling, L., Fernandez-Villaverde, J., Uhlig, H., 2024. Central bank digital currency: When price and bank stability collide. J. Monetary Econ. 145.
Sprague, O.M.W., 1910. History of Crises Under the National Banking System. US Government Printing Office.
Voellmy, L., 2024. Decomposing liquidity risk in banking models. SNB Working Paper 2024-03.
Wallace, N., 1990. A banking model in which partial suspension is best. Fed. Reserve Bank St. Louis Q. Rev. 14 (4), 11–23.
Wicker, E., 2000. Banking Panics of the Gilded Age. Cambridge University Press.
Williamson, S., 2012. Liquidity, monetary policy, and the financial crisis: A new monetarist approach. Amer. Econ. Rev. 102 (6), 2570–2605.
World Bank, 2021. The Global Findex Database 2021: Financial Inclusion, Digital Payments, and Resilience in the Age of COVID-19.