Discussion Paper Series – CRC TR 224

Discussion Paper No. 605
Project C 03

Optimal Banking Arrangements: Liquidity Creation Without Financial Fragility

Maxi Günnewig [1]
Yuliyan Mitkov [2]

October 2024

[1] Department of Economics, University of Bonn, Email: mguennewig@uni-bonn.de
[2] Department of Economics, University of Bonn, Email: ymitkov@uni-bonn.de

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Optimal Banking Arrangements: Liquidity Creation without Financial Fragility*

Maxi Guennewig[†] Yuliyan Mitkov[†]

October 16, 2024

Abstract

Diamond and Rajan (2000, 2001) argue that banks create liquidity by issuing deposits to fund difficult, illiquid firms that otherwise cannot obtain funding. Since deposits may lead to bank runs, this resulting financial fragility is *essential* for liquidity creation. We revisit the Diamond-Rajan model of financial intermediation and show that a bank with an optimal financing structure is not subject to runs. Our contract rests on three simple notions. First, each bank creditor has the right to demand repayment at every instant. Second, the repayment is given by the value of a pre-specified fraction of the bank's assets. Third, some creditors are more senior than others: their repayment demands are prioritized. In contrast to Diamond and Rajan, we find that financial fragility is *detrimental* to liquidity creation.

Keywords: Liquidity, banking, financial fragility, optimal contracts, collateral.

JEL Codes: G21

*We thank Charles Calomiris, Alkis Georgiadis-Harris, Todd Keister, and David Skeie for their helpful comments. The authors gratefully acknowledge support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the CRC TR 224 (Project C03).

[†]Department of Economics, University of Bonn. Adenauerallee 24-42, 53113 Bonn, Germany.

1 Introduction

Banks issue deposits, which allow their creditors to withdraw their funds when needed. They also issue loans, which channel these funds to borrowers seeking to finance profitable investments. Each of these activities is valuable on its own. However, when combined, banks become fragile. Indeed, a loss on the asset side or even a rumor thereof can trigger a run, leading to the bank's failure. But what compels the same institution to issue both loans and deposits?

One popular view is that this practice originated from historical precedent and was subsequently solidified by government guarantees (Kareken, 1985; Pennacchi, 2012). In an influential body of work, Diamond and Rajan (2000, 2001) put forward a different logic. They argue that banks issue deposits precisely because it allows them to fund difficult, illiquid borrowers who cannot obtain funding from equity-financed banks or from financial markets directly. In doing so, banks *create liquidity* for such borrowers. Since issuing deposits leads to bank runs, this resulting financial fragility is *essential* for liquidity creation.

In this paper, we revisit the Diamond and Rajan (2000, 2001) model of financial intermediation. We study the bank's contracting problem and show that banks with an optimal financing structure are not subject to runs. Thus, and in contrast to Diamond and Rajan, we find that financial fragility is not essential but *detrimental* to liquidity creation.

Central to the model is limited commitment. Entrepreneurs have access to a profitable project but need funding. As in Hart and Moore (1994), they can commit their human capital to the project only on a spot basis. A lender who has issued a loan to the entrepreneur can extract future repayment by threatening to liquidate the project. That is, she can take the project's physical assets from the entrepreneur and put them to alternative use. However, since the entrepreneur is the first-best user of the project's assets but cannot commit his future human capital, the entrepreneur cannot commit to fully repaying the project's cash flows. Instead, the lender can extract only part of the project's cash flows when collecting the loan. Because of this limited commitment, the entrepreneurs' projects cannot be sold or borrowed against for the full cash flow they are expected to generate.

An early-stage lender (the banker henceforth) knows how the project is set up and learns the best way of redeploying the project's physical assets, thus increasing the project's liquidation value. This superior liquidation threat gives the banker specific loan collection skills, as it allows her to extract a greater repayment from the entrepreneurs than any other creditor. Nevertheless, illiquidity remains an issue: after making the loan, the banker may need to sell it or borrow against it to finance consumption needs or undertake new investments.[1] However, the banker cannot commit to using her specific loan collection skills on behalf of her own creditors in the future. Instead, the banker seeks to renegotiate with her creditors to extract more rents, asking to be compensated for using her specific skills. It follows that the banker can borrow less against the loan than the repayment she can collect from the entrepreneur. The loan is poor collateral.

Consequently, both the entrepreneurs' projects and the loan to the entrepreneur cannot be financed to the full extent of their expected cash flows: they are *illiquid*. This illiquidity can be socially costs when it affects the initial terms of the loans to the entrepreneurs. If the banker cannot borrow against the loan portfolio when she needs to, she may reduce her activity and finance fewer socially desirable projects, thus impairing the flow of credit in the economy.

Diamond and Rajan (2001) propose an elegant solution to the illiquidity problem: the banker issues demand deposits. Deposits have three key design features. First, each depositor can present his claim at any time and demand repayment. Second, this repayment is given by the pre-specified (and thus fixed) face value of each claim. Third, depositors are serviced sequentially. That is, they are repaid on a first-come, first-served basis. As a consequence, each depositor is repaid in full whenever he demands it—unless too many other depositors have done so beforehand and the bank has run out of funds.

Any attempt on the banker's side to seek rents and repay less than originally promised precipitates a run: depositors rush to the bank to fully recover their claims before others do so. Such behavior is individually rational for each depositor, even if a bank run is costly to the group of depositors. In other words, the deposit contract creates a collective action problem that commits the depositors to not making any concessions. A run fully disintermediates the banker and drives her rents to zero. Consequently, she prefers to avoid them by repaying what was promised to the depositors in full. Thus, issuing demand deposits can be a commitment device for the banker, tying her specific skills to the loan portfolio. The loan portfolio can then be collateralized to the full extent of its expected cash flows: it becomes liquid.

However, demand deposits have two major drawbacks. First, the deposits' first-come, first-served rule leads to equilibrium multiplicity. Indeed, the bank may experience a panic

[1]Modeling such liquidity shocks, which force the banker to borrow after making the loan, is not the only way to set up the problem (see Diamond and Rajan (2001)). Indeed, the results are the same if each potential lender has insufficient funds to finance the entrepreneur on her own and, therefore, must raise external funds. If educating the initial lender is costly, delegating this task to one lender (i.e., the banker) is efficient. The banker then again lacks commitment to use her loan collection skills on behalf of her own creditors.

run, as in Diamond and Dybvig (1983). Second, the bank experiences fundamental runs if the outstanding deposits exceed the project's cash flows even without opportunistic behavior by the banker. The collective action problem among depositors stands in the way of any useful renegotiations that adjust the terms of the deposit contract in response to real asset losses.

Diamond and Rajan (2000) show that, for the reasons above, banks may wish to complement deposits with a softer claim referred to as 'capital.' One interpretation of capital is bank equity, which grants its holders the right to replace the banker at any time.[2] Capital holders are junior to depositors. Furthermore, they are not serviced sequentially and, therefore, do not suffer from the collective action problem that gives rise to runs. As a consequence, capital can be renegotiated downwards and serve as a buffer against adverse shocks to the value of the bank's assets. The downside is that it allows rent-seeking in the absence of adverse shocks—the illiquidity problem is back. Thus, greater bank capital reduces financial fragility but also hampers liquidity creation.

We show that fragility can be eliminated without undermining liquidity creation. Specifically, we design a contract that deters rent-seeking, facilitates useful renegotiations in response to losses, and is not subject to coordination failures.

Our contract has three key design features. First, each bank creditor can present his claim at any time and demand repayment. Second, each creditor's repayment is given by the value of a pre-specified *share* of the bank's assets. Third, some creditors are designated as *senior*: they are repaid in full when the bank cannot service all creditors that demand repayment.

Our contract thus shares the first feature with the deposit contract, but it differs regarding the other two. In particular, the deposit contract specifies a promised repayment of $\$D$. Our contract specifies a promised repayment of $s\%$ of bank assets. Furthermore, given the contractually specified priority provisions, it induces 'orderly' withdrawal: the senior creditors always withdraw first. On the other hand, sequential service under the deposit contract induces 'disorderly' withdrawal: even though all depositors hold the same claim, some are lucky and withdraw first, while others may be unlucky and cannot withdraw.

Under the optimal contract, the bank creditors demand immediate repayment whenever the banker seeks rents by offering to repay less than the entire cash flow of the loan portfolio. In this case, the banker is disintermediated and her rents are driven to zero. Importantly,

[2] Another interpretation is long-term debt, which grants creditors this right only if the banker defaults. An advantage of equity over long-term debt is that some events meriting replacing the banker will not trigger default since they are not verifiable.

since investors withdraw shares of the banker's assets, their option value of withdrawing naturally fluctuates with the state of the economy and perfectly adjusts to the returns of the loan portfolio. This allows our contract to accommodate real asset risk. At the same time, all attempts to renegotiate the investors' claims beyond adjustments to the state of the economy are bound to fail. The optimal contract, therefore, preserves the desired property of the deposit contract—the banker cannot seek rents—but is not subject to fundamental runs.

Furthermore, the priority provisions among bank creditors prevent coordination failures that can lead to bank runs even without rent-seeking. The usual logic around runs—everybody rushing to the bank to get their money out before others do so—does not apply. Indeed, the junior creditor can never 'outrun' the senior creditor. If the junior creditors are not incentivized to demand immediate repayment, there is no reason for the senior creditors to do so. The run equilibrium unravels. It follows that the banker is only disintermediated when she is rent-seeking—and not due to coordination failures and low asset returns. The trade-off between liquidity creation and fragility disappears.

Importantly, the banker is free to issue claims other than deposits if her primary goal is to enhance the collateral value of her assets. Nothing in the set-up of Diamond and Rajan (2000, 2001) forces the banker to serve withdrawal requests sequentially rather than according to some pre-specified rule. In other words, sequential service is a design choice and not a constraint.[3] This is demonstrated by the banker's ability to issue both demand deposits and capital (which is not subject to sequential service) at the same time. Indeed, we show that sequential service is not part of the optimal contract since it generates excessive fragility. Moreover, our optimal contract highlights that there is no role for capital to act as a buffer against adverse shocks to the value of bank assets.

Finally, we stress that our optimal contract is consistent with bank creditors' role as bank monitors. Calomiris and Kahn (1991) argue that banks issue demand deposits as their first-come, first-served rule sets the correct incentives for depositors with high monitoring capabilities. The idea is that the monitoring depositors can successfully withdraw from the bank at the first sign of trouble. This, in turn, alerts the remaining uninformed creditors that they should also withdraw, and the banker is deterred from misbehaving. The monitoring incentives under the optimal contract outlined in our paper remain intact if those creditors with high monitoring abilities are prioritized (i.e., they hold the senior claims). The optimal

[3]Diamond and Rajan (2005) explicitly write that "the sequential service constraint inherent in demand deposits, which is the source of bank runs and systemic fragility, is not superimposed in our framework but is necessary for the function the banks perform." (p. 618)

contract further avoids the cost of coordination failures associated with demand deposits.

Contribution to the literature. Our paper is primarily related to two strands of literature: financial intermediation and creditor dispersion. We begin by discussing the former.

Previous work has studied how financial intermediaries can efficiently allocate liquidity innate to certain assets among different economic agents. For example, intermediaries offer diversification across depositors (Diamond and Dybvig, 1983), borrowers (Holmström and Tirole, 1998), and even banks (Bhattacharya and Gale, 1987; Allen and Gale, 2000). In contrast, Diamond and Rajan (2000, 2001) argue that banks can enhance the liquidity of their assets by issuing first-come, first-served deposits. Choosing a fragile liability structure is thus inevitable for liquidity creation. Our contribution is to show that first-come, first-served demand deposits are not optimal in their framework. In contrast, the optimal contract is not subject to runs. As a consequence, fragility is not inevitable but detrimental to liquidity creation.

Our paper also relates to theories of why creditor dispersion may be beneficial (e.g., Berglöf and Von Thadden (1994); Bolton and Scharfstein (1996)). If a single lender can commit to liquidating if borrowers attempt renegotiation (as in Diamond (1984)), then borrowers are indeed deterred from renegotiating. Multiple lenders will not be required to discipline borrowers. But what if lenders cannot commit to liquidating? Borrowing from multiple lenders makes individual lenders more eager to liquidate if it imposes a negative externality on other lenders. At the same time, having multiple lenders may lead to collective action problems (Diamond and Rajan, 2000, 2001).[4] Our contribution is to show that this need not occur under the optimal contract.

Furthermore, Diamond and Rajan (2000, 2001) provide a theory of financial intermediation: demand deposits and the contract we derive in Section 3 disciplines banks but not industrial firms. It should be noted that a fragile liability structure can be a commitment device for firms (see, e.g., Von Thadden et al. (2010)). Specifically, if a firm's asset can be quickly and irreversibly liquidated in a run to repay debt holders, the threat of a run on the firm's assets commits the firm to repay more than the liquidation value of its assets. In those cases, our contract is useful to all borrowers. However, liquidation is usually inefficient, implying that the parties would attempt to negotiate around it. When the decision to liquidate, sell assets, or renegotiate with the borrower is made after the run, a demandable

[4]This logic goes back to the theory of the second best (Lipsey and Lancaster, 1956). Specifically, if removing one distortion (e.g., limited commitment) is infeasible, then introducing another distortion (collective action problem) may partially or fully offset the first distortion and lead to more efficient outcomes.

debt structure may fail to discipline firms.[5] We elaborate on why this is the case in Section 8.

In Calomiris and Kahn (1991), demand deposits subject to sequential service also commit borrowers to repay more.[6] There are two important differences. First, in their model, demandable debt holders force liquidation of the borrower's project in some states of the world. In doing so, they prevent a 'crime in progress,' e.g. a diversion of funds. This agency friction differs from the one in our setup, where the banker cannot divert funds but instead threatens not to use her skills on behalf of her creditors. Second, the setup in Calomiris and Kahn (1991) implies that bank runs are part of the optimal contract between the bank and its depositors. In other words, bank runs are necessary to prune the financial system of bankers not acting in the depositors' best interest and, thus, to achieve efficiency. In contrast, bank runs do not have any useful function in the Diamond-Rajan framework.

Outline. The rest of the paper is organized as follows. Section 2 introduces a simple framework we use to establish our main result. Section 3 proves our main result, while Section 4 contains a simple illustration to provide the intuition. Section 5 revisits the demand deposit contract (with some capital buffers) as in Diamond-Rajan and explains why it is not optimal. We discuss our results and their implications in Section 6. Section 7 studies two extensions of the baseline setup. The full model of financial intermediation is outlined in Section 8. Finally, Section 9 concludes.

2 Model

2.1 Environment

Our baseline model contains the core element of the models in Diamond and Rajan (2000, 2001; henceforth Diamond-Rajan). The model has three dates and includes a banker and many investors. All agents have linear utility over consumption in all periods and do not

[5]Some authors interpret debt collection as taking the borrower to lengthy and inefficient court procedures (e.g., Diamond (2004)). In that case, demandable debt disciplines all issuers. Dispersed creditors might commit the borrower to repay for other reasons, e.g., by making renegotiation or default more costly for the borrower (Ivashina et al., 2016). Alternatively, dispersed creditors might collect a lower overall repayment but, due to free-riding, also engage in less socially wasteful rent-seeking (Bris and Welch, 2005).

[6]Calomiris, Kahn, and Krasa (1991) show how the model is easily extended to be a model of financial intermediation.

discount future payoffs. The banker receives a unit endowment on date 0 and has the option to invest this endowment in a loan portfolio consisting of a continuum of identical loans.[7]

The state of the economy θ is realized on date 2. The CDF of the state is $F(.)$ with support $[\underline{\theta}, \overline{\theta}]$. The loan portfolio yields a state-dependent cash flow of $x(\theta) \in (0, \infty)$ if the banker deploys her loan collection skills. The loan portfolio's expected cash flow is thus $\mathbb{E}[x(\theta)] = \int_{\underline{\theta}}^{\overline{\theta}} x(\theta) dF(\theta)$. We assume that cash flows are observable and verifiable, and the banker cannot divert them. At the same time, the state of the economy θ is observable but not verifiable. The agents thus cannot write contracts directly contingent on the realized state but can write contracts contingent on the realized cash flow of the loan portfolio. There is also a storage opportunity with a gross return of one available to all agents.

The banker may experience a liquidity shock at date 1—and so after she has obtained the loans—with probability $q \in [0, 1]$. As in Diamond and Rajan (2001), the liquidity shock may arise due to a highly valued investment opportunity at date 1 or due to a shock to the banker's discount rate. In case of such a shock, the banker wants to borrow as much as possible, using the loans as collateral. Each investor is deep-pocketed on dates 1 and 2. Thus, there is no aggregate shortage of liquidity. The banker signs a contract with one investor or a group of investors that specifies some repayment $B_i(x)$ to each investor i as a function of the loan portfolio's cash flow x on date 2.

We define a *default* as the banker missing a contractually specified payment or withholding her loan collection skills. In case of default, the contract may grant the investors the right to seize (a portion of) the banker's assets, namely the loan portfolio.[8] We say that the banker is *disintermediated* if the investors have seized all her assets. This implies that the ownership over said loans is transferred from the banker to the investor, who now collects the repayment of the loans. However, each investor lacks the banker's specific skills and can only collect a repayment of $\beta x(\theta)$ from the portfolio of loans in state θ, where $0 < \beta < 1$. This implies a market value of the loans of $\beta\mathbb{E}[x(\theta)]$ at date 1 and of $\beta x(\theta)$ at date 2.

The central friction in this setup is limited commitment: at any date t the banker can commit to using her specific skills only for that date. Thus, she cannot commit at date 1 to using her specific skills when collecting the loans on behalf of investors at date 2. In other words, the banker has intra-temporal but not intertemporal commitment. The justification for this assumption is as in Hart and Moore (1994): the law prevents agents from irrevocably

[7]In the full model (Section 8), entrepreneurs can invest in a project but have no endowment. The banker then issues loans to entrepreneurs.

[8]Alternatively, the contract may grant non-contingent rights to seize assets, which is useful when not all default events can be verified.

selling themselves into bondage. Thus, courts cannot compel the banker to deploy her human capital even if they can enforce other aspects of contracts. Consequently, on date 2, the banker may seek to renegotiate financial contracts signed on date 1, as explained shortly.

We assume the following. First, the loan portfolio has a positive net present value, $\mathbb{E}[x(\theta)] > 1$. Second, it has a negative net present value if it is collected by investors without specific skills, $\beta\mathbb{E}[x(\theta)] < 1$. Third, the loan portfolio is perfectly divisible: suppose the banker retains ownership of a fraction of $1 - s$ of the loans, whereas an investor has seized the remaining fraction of s. Then, if the state of the economy is θ, the banker can collect a cash flow of $sx(\theta)$, whereas the investor can only collect $(1 - s)\beta x(\theta)$.

Figure 1 summarizes the sequence of events.

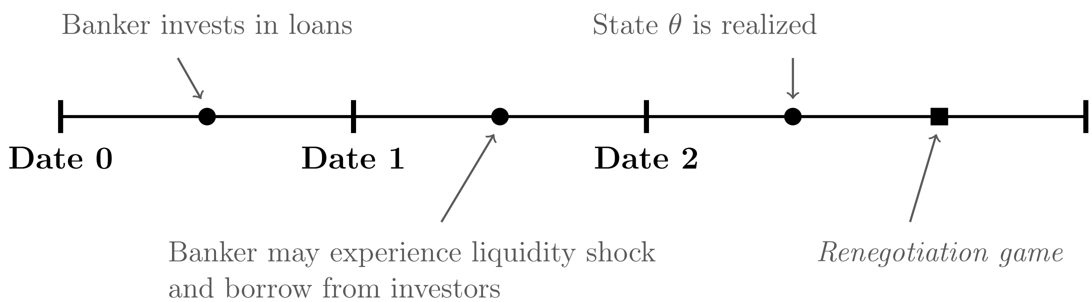

Figure 1: Timeline of events.

Renegotiation. After the state θ is realized, the banker and the investors play a non-cooperative game, which we call the *renegotiation game*. The banker moves first by offering each investor a new repayment. The investors move second, each simultaneously and independently deciding whether to accept or reject the banker's offer.

Suppose the cash flow if the banker deploys her loan collection skills is x. We follow Diamond-Rajan in assuming that the banker makes perfectly observable take-it-or-leave-it offers to each investor i of the following kind: "I will not deploy my loan collection skills unless you agree to be paid $\tilde{B}_i(x) \in [0, x]$." Assuming that the banker has full bargaining power entails no loss of generality. Key to our findings is that the banker has *some* bargaining power and thus seeks rents at date 2.

If investor i accepts, he now holds a claim of $\tilde{B}_i(x)$ on the banker. If investor i rejects, he seizes a portion of the banker's assets. If multiple investors reject the banker's offer and the respective rights to seize assets become incompatible (the individual shares sum to more than one), then the date-1 contract determines the allocation of assets among the investors. In other words, the investors' outside option when renegotiating with the banker is a function of the other investors' actions and the contractually specified *rationing mechanism* reconciling incompatible claims on the banker's assets.[9] We provide example contracts highlighting such rationing rules in the following subsection.

The banker's renegotiation offer and the rationing mechanism in the original contract define a subgame. Our solution concept is subgame perfection. We interpret offering $\tilde{B}_i(x) = B_i(x)$ as the banker not renegotiating with investor i and insist on the feasibility of the banker's offers: $\sum_{i \geq 1} \tilde{B}_i(x) \leq x$. The banker can also impose an infinitesimal fee on investors who reject her renegotiation offer. The fee is specified in the initial contract and collected when an investor seizes assets. The purpose of the fee is to induce those indifferent between accepting and rejecting the banker's offer to accept.[10]

Notice that an investor's payoff depends not only on his action to accept or reject the banker's offer but also on the other investors' actions. This strategic interaction between the investors occupies a central place in the analysis. Also, while the banker is committed to using her specific skills to repay $\tilde{B}_i(x)$ after investor i has accepted, the potential (partial) seizure of the banker's assets by other investors may stand in the way of the full repayment of investor i's claim.

Finally, the banker's payoff may be zero because she got fully disintermediated or because she transferred all of the loan repayments to the investors. We assume the banker strictly prefers the latter to the former. This assumption can be justified through a (small) non-pecuniary benefit $b > 0$ of retaining ownership of the loan portfolio (e.g., reputational concerns).

This concludes the description of the renegotiation game. Figure 2 summarizes the sequence of events within the game.[11]

[9] Von Thadden et al. (2010) distinguishes between individual and collective debt collection, with the latter interpreted as bankruptcy. Legal scholars have long viewed bankruptcy as adjusting (i.e., pairing down) individual claims when these claims are inconsistent (Jackson, 1986). As in Von Thadden et al. (2010), we analyze a model where collective debt collection rules are a part of the original contract.

[10] Diamond-Rajan follows a similar approach by (implicitly) stipulating that investors indifferent between accepting and rejecting the banker's offer will accept. Also, instead of a contractually specified fee, we can assume that the investors incur a small inconvenience cost for seizing assets.

[11] Our renegotiation game is inspired by Diamond-Rajan (Diamond and Rajan (2001), p. 309). The key difference is that we do not restrict the rationing mechanism to sequential service.

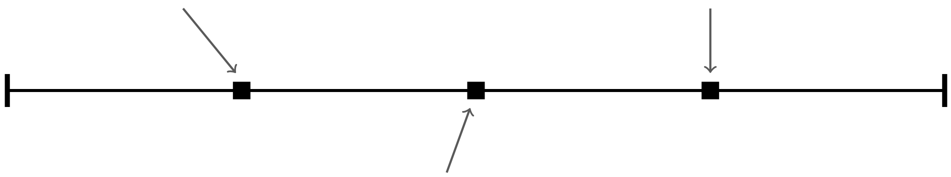

Figure 2: Sequence of events within the renegotiation game.

Remarks. In the full framework (described in Section 8), the banker becomes a banker by intermediating funds between many investors and an entrepreneur who has access to a productive project but lacks the necessary funds to invest at date 0. The banker then seeks to render the loan to the entrepreneur liquid by choosing an optimal financing structure, which we describe in the next section. We focus on the simpler version since augmenting the model with these extra components does not alter our results.

2.2 Example contracts

In this subsection, we describe three example contracts.

The first contract is with a single investor, who is promised a repayment of $B(x)$ if the loan portfolio's cash flows on date 2 are x, where $B(x) \in [0, x]$. This investor also has the right to seize assets should the banker default on the claim at date 2. The rationing rule is immaterial in this case. A debt contract with a face value of $D > 0$ is a special case corresponding to $B(x) = \min\{x, D\}$.

The second contract is the deposit contract issued to $n \geq 2$ investors, as in Diamond-Rajan. Each depositor is promised a repayment of $D/n > 0$ on date 2. Depositors can demand repayment by seizing loans at every instant, even if the banker is not renegotiating or defaulting on their claim. If a depositor chooses to withdraw, he receives units of the underlying loans with a market value of D/n—unless too many depositors demand repayment simultaneously so that the banker cannot repay all of them. If so, the rationing rule is *sequential service*. Withdrawing depositors are serviced one by one in a randomly determined order. The first depositors in that order are repaid in full until all of the banker's assets have been seized; the remaining depositors then seize no assets.

The third contract is a simplified version of the optimal contract described in the following section. The banker borrows from two investors $i \in \{1, 2\}$. Each investor i is promised a repayment of $D_i > 0$ and receives the right to seize a portion $s_i \in [0, 1]$ of the loan portfolio should the banker default. For example, investor 1 can seize 60% of the assets while investor 2 can seize 50% of the assets. If so, it cannot be that both seize their respective portions. The rationing rule is then as follows. Investor 1 is more senior than investor 2. Consequently, if both investors want to seize assets, then investor 1 seizes their promised share of 60% while investor 2 can only seize the remaining assets. If only investor 2 seizes assets, then he seizes his promised portion of 50%.

2.3 Source of illiquidity

The loans' cash flows can only be fully collected using the banker's specific skills. Since, as we will see, these cannot be easily committed, the loan portfolio is *illiquid*—only part of the expected cash flows can be pledged to the investors.

We begin by illustrating a financing arrangement that fails to render the loan portfolio liquid. Suppose the state of the economy is θ, implying that the loan portfolio yields a cash flow of $x(\theta)$ should the banker deploy her loan collection skills. Suppose further that the banker has borrowed from one investor and promised to repay $B(x(\theta))$. If $B(x(\theta)) > \beta x(\theta)$, the banker makes the following offer: "I will not deploy my loan collection skills unless you agree to be paid $\beta x(\theta)$."

If the investor accepts, the banker uses her loan collection skills. The banker's payoff is $(1 - \beta)x(\theta)$, whereas the investor's payoff is $\beta x(\theta)$. If the investor rejects, he seizes the loan portfolio and collects $\beta x(\theta)$. Thus, the investor is indifferent between accepting and rejecting the banker's offer and accepts. The maximum expected payment the banker can pledge to this single investor is thus $\beta \mathbb{E}[x(\theta)]$ even though the expected loan repayment is $\mathbb{E}[x(\theta)]$. In other words, the loans are illiquid.

Consequently, the banker may prefer not to obtain the loan portfolio in the first place. To illustrate, suppose that the banker experiences the liquidity shock at the interim date with probability one. Purchasing the loan portfolio at date 0 and using it as collateral to borrow from one investor at date 1 using the financial arrangement above yields a return of $\beta \mathbb{E}[x(\theta)] < 1$. The banker then prefers to store her endowment as storage yields a return of one.

3 Main result

3.1 Optimal contract

We show how to design a contract that renders the loan portfolio fully liquid in every state of the economy. The banker borrows from $n \geq 2$ investors on date 1 via the contract C^* specifying promised repayments, rights to seize assets, and a rationing mechanism in case the investors want to seize more than the available assets. Specifying the promised repayments is unnecessary: What matters is the investor's right to seize the banker's assets. Specifically, irrespective of whether or not the banker has defaulted, investor 1 has the right to seize all of the assets. Each of the other $n-1$ investors has the right to seize a share $\frac{1-\beta}{\beta(n-1)}$ of the assets. That is,

$$s_1 = 1 \qquad \text{and} \qquad s_i = \tfrac{1-\beta}{\beta(n-1)} \qquad \text{for } i = 2, ..., n.$$

Suppose the state is such that the loan portfolio generates a cash flow of x should the banker deploy her skills. By seizing their portion of the loans, investor 1 can thus generate a payoff of βx; any other investor can generate a payoff of at most $\frac{1-\beta}{(n-1)}x$.[12] The contract includes a vanishing fee incurred by each investor when they seize assets.

Notice also that not all investors can exercise their rights to seize assets simultaneously, as their shares sum to more than one. The rationing mechanism in that case is simple: investor 1 is *senior* to all other investors. If multiple investors, including investor 1, attempt to seize assets, then investor 1 seizes all assets, and consequently, all other investors seize zero assets. The seniority arrangement among the junior $n-1$ investors is irrelevant. Hence, we leave it unspecified.

We are now ready to state our main result.

Theorem. *Suppose the banker issues contract C^* on date 1. If n is sufficiently large, the banker is not disintermediated and repays a total of $x(\theta)$ to the investors in every equilibrium of the renegotiation game and in every state $\theta \in [\underline{\theta}, \overline{\theta}]$. The contract C^* thus allows the banker to borrow $\mathbb{E}[x(\theta)]$.*

We prove the theorem in the following subsection. We explain the intuition using a two-investor illustration in Section 4.

[12]Notice that investors have non-contingent rights to seize assets. This is not crucial. Section 6.1 provides alternative specifications of the optimal contract.

3.2 Proof

Suppose the state of the economy is such that the loan portfolio's cash flows are given by x should the banker deploy her skills. Let $\tilde{B}_i(x)$ denote the banker's offer to investor i. First, note that each offer $\left(\tilde{B}_i(x)\right)_{i=1}^{N}$ induces a subgame with finite players with a finite set of strategies; hence, a Nash equilibrium exists for each offer (see, e.g., Fudenberg and Tirole (1991)).

We base the proof on three lemmas. The first lemma shows that no equilibrium exists in which some junior investor(s) reject and the senior investor accepts. The second lemma highlights that the banker is disintermediated unless she offers βx to the senior investor and $\frac{1-\beta}{n-1}x$ to each junior investor. The final lemma shows that all investors accept if the banker offers βx to the senior investor and $\frac{1-\beta}{n-1}x$ to each junior investor. The total payment to the investors is x.

For the remainder of the proof, we set $\beta > \frac{1-\beta}{\beta(n-1)}$, which holds for n sufficiently large. Throughout the proof, we treat the fee for seizing assets κ as strictly positive but arbitrarily close to zero and omit it from the notation except in the proof of Lemma 2.

Lemma 1. *There exists no equilibrium in which $k \in \{1, 2, ..., n-1\}$ junior investors reject, and the senior investor, as well as the remaining $(n-1-k)$ junior investors, accept.*

Proof. Consider a candidate equilibrium in which $k \in \{1, 2, ..., n-1\}$ junior investors reject, and the remaining junior and senior investors accept. We now rule out this candidate equilibrium by showing that at least one investor has a strictly profitable unilateral deviation to reject. First, the banker has $S(k)$ shares of the loan portfolio left after k junior investors have seized assets, where

$$S(k) \equiv \max\left\{1 - k\tfrac{1-\beta}{\beta(n-1)},\, 0\right\}.$$

We now argue that in this candidate equilibrium it must be that $S(k) < \frac{1-\beta}{\beta(n-1)}$. To this end, suppose otherwise: $S(k) \geq \frac{1-\beta}{\beta(n-1)}$. Hence, each of the $n-1-k$ junior investors who accept can instead seize their share of the assets in full and get a payoff of $\frac{1-\beta}{(n-1)}x$. Consequently, each junior investor who accepts must be paid at least $\frac{1-\beta}{(n-1)}x$. Also, the senior investor must be paid at least βx since he can always seize all assets by rejecting and thus getting βx. Hence, the total payment to the $n-k$ investors who accept must be no less than

$$\beta x + (n-1-k)\frac{1-\beta}{(n-1)}x.$$

The total repayment to the accepting investors is bound from above by $S(k)x$. Hence, for this candidate equilibrium to be indeed an equilibrium, we must have

$$\left(1 - k\tfrac{1-\beta}{\beta(n-1)}\right) x \;\geq\; \beta x + (n-1-k)\frac{1-\beta}{(n-1)}x.$$

Simple algebraic manipulation reveals a contradiction if $\beta \in (0,1)$. Hence $S(k) < \frac{1-\beta}{\beta(n-1)}$, which implies that the senior investor's payoff from accepting is bound from above by $\frac{1-\beta}{\beta(n-1)}x$. Since $\beta > \frac{1-\beta}{\beta(n-1)}$, the senior investor is strictly better off rejecting the banker's offer. The claim follows. □

For the next lemmas, define $B_1^*(x) = \beta x$ and $B_i^*(x) = \frac{1-\beta}{n-1}x$ for all $i \geq 2$. Thus, $B_i^*(x)$ corresponds to the payoff investor i can generate by rejecting the banker's offer and seizing loans if all other investors accept. Also, recall that the banker gets disintermediated if the investors seize all assets. Next, recall that the banker's payoff when she averts disintermediation is at least equal to the non-pecuniary benefit of retaining ownership of the assets b, where this benefit can be arbitrarily small.

Lemma 2. *Suppose $\tilde{B}_i(x) < B_i^*(x)$ for some i. Then, the banker is disintermediated.*

The proof is in Appendix A, where we show that the banker is disintermediated with a probability that goes to one as the fee for seizing assets vanishes. The intuition is straightforward: the senior investor never accepts an offer less than $B_1^*(x) = \beta x$ since he can always seize all assets and get a payoff of $B_1^*(x)$. To avoid disintermediation, the banker must offer at least $B_1^*(x)$ to the senior investor. If the senior investor accepts, then any junior investor i will not accept an offer less than $B_i^*(x) = \frac{1-\beta}{n-1}x$. Finally, as a consequence of Lemma 1, the senior investor will not accept the banker's offer unless all junior investors accept it. Hence, the banker can induce all investors to accept only by offering $B_i^*(x)$ to each investor.

Lemma 3. *Suppose $\tilde{B}_i(x) = B_i^*(x)$ for all i. The subgame induced by this offer has a unique equilibrium in which all investors accept.*

Proof. Let $\tilde{B}_i(x) = B_i^*(x)$ for all i. Suppose all investors accept the banker's offer. By rejecting, investor i gets $B_i^*(x)$ whereas by accepting he gets $\tilde{B}_i(x)$. Each investor accepting is thus an equilibrium of this subgame. Moreover, this is the only equilibrium. In particular, recall that the contract C^* imposes an infinitesimal rejection fee, making each investor accept when indifferent between accepting and rejecting. By Lemma 1, the senior investor strictly prefers to reject whenever at least one junior investor rejects. If the senior investor rejects, he seizes all assets, and all other investors get a payoff of zero regardless of whether they accept or reject. Hence, they accept. But if all junior investors accept, the senior investor will be indifferent between accepting and rejecting, as he achieves a payoff of $B_1^*(x)$ either way. Thus, he accepts. □

By combining the three lemmas, we conclude that the banker is disintermediated unless she offers βx to the senior investor and $\frac{1-\beta}{n-1}x$ to each of the $n-1$ junior investors for a total repayment of x. As a consequence, the senior investor is willing to lend $\beta\mathbb{E}[x]$ to the banker at date-1, and the junior investors are willing to lend a total of $(1-\beta)\mathbb{E}[x]$. This completes the proof.

4 An illustration with two investors

We now explain the intuition behind our result by illustrating the optimal contract with just two investors. For this subsection only, set $\beta \in \left(\frac{1}{2}, 1\right)$.[13]

Suppose the banker has issued contract C^* to two investors at date 1, with investor 1 being senior and investor 2 being junior. Suppose the realized state of the economy is such that the loan portfolio generates a cash flow of x on date 2. The banker begins renegotiations by offering a payment of $\tilde{B}_1(x)$ to the senior investor and $\tilde{B}_2(x)$ to the junior investor. Recall that the senior investor has the right to seize all of the bank's assets should he reject the banker's offer; the banker is then disintermediated. The junior investor can seize a share $\frac{1-\beta}{\beta}$ of the bank's assets, but only if the senior investor accepts the banker's offer. Table 1 shows the payoffs of each investor for all four strategy combinations.

	Junior investor accepts	Junior investor rejects
Senior investor accepts	$\tilde{B}_1(x)$, $\tilde{B}_2(x)$	$(1-\frac{1-\beta}{\beta})x$, $(1-\beta)x$
Senior investor rejects	βx, 0	βx, 0

Table 1: Payoff matrix given the banker's offer.

The senior investor can always guarantee a payoff of βx by seizing all assets. Consequently, this investor rejects any offer $\tilde{B}_1(x) < \beta x$. The banker is disintermediated, and her payoff is zero. Consider therefore an offer $\tilde{B}_1(x) = \beta x$ to the senior investor. Feasibility then

[13] A larger number of investors is only needed if $\beta \leq \frac{1}{2}$.

requires that $\tilde{B}_2(x) \leq (1-\beta)x$. We now argue that any offer $\tilde{B}_2(x) < (1-\beta)x$ to the junior investor is not accepted, and the banker is disintermediated as a consequence.

First, suppose $\tilde{B}_2(x) < (1-\beta)x$. If the senior investor accepts the banker's offer, the junior investor strictly prefers rejecting the offer to seize assets worth $(1-\beta)x$. The senior investor then finds it strictly profitable to also reject. Given his priority, he seizes all assets. The banker is disintermediated.

Second, suppose $\tilde{B}_2(x) = (1-\beta)x$. There is an equilibrium where both investors accept the banker's offer. As a consequence, the banker transfers the entire cash flow of the loan portfolio. There is another equilibrium where both investors reject. However, the second equilibrium is unstable: the junior investor is indifferent between accepting and rejecting since he gets a zero payoff regardless of his action.

In the proof of the Theorem, we show how an infinitesimal rejection fee unravels the rejection equilibrium.[14] The intuition is that the small fee acts as a tie-breaker, inducing the junior investor to accept. The senior investor then also accepts. It follows that an offer to transfer the entire cash flow of the loan portfolio is always accepted in every state of the economy. In other words, there are no runs following such an offer, neither due to bad bank fundamentals nor due to investor panics.

Since the banker prefers not to be disintermediated, she transfers the loan portfolio's total cash flows x to the investors. Consequently, the banker can borrow $\mathbb{E}[x(\theta)]$ using the loan portfolio as collateral. The loan portfolio is fully liquid.

Intuition. The intuition behind our result is the following. On the one hand, rent-seeking is harmful as it reduces the collateral value of the bank's assets and, thus, liquidity creation. On the other hand, renegotiations are useful as they allow the adjustment of contractual terms to the state of the economy ex-post. Our contract balances the two. It specifies that the investors can seize portions of the bank's assets whose market value fluctuates with economic conditions. Hence, the value of the investors' option to foreclose on the bank naturally adjusts to the state of the economy. This feature facilitates useful renegotiations in bad states of the world and allows our contract to accommodate real asset risk.

At the same time, the banker cannot extract rents: all attempts at renegotiations beyond

[14] Recall: the fee is specified in the initial contract and paid to the banker as the investors seize assets. Alternatively, the fee may stem from some small inconvenience of seizing assets, e.g., filling out forms. We should note that the most straightforward approach (with the same implications) is to impose a tie-breaking restriction on investors' strategies: each investor accepts when indifferent. Unfortunately, in that case, equilibrium existence is not guaranteed for all possible renegotiation offers.

adjustments to the state of the economy are bound to fail. Each investor rejects a repayment offer that is less than the value of the banker's asset, which he can seize. Whenever all investors accept the banker's repayment offer, the sum of the payoffs from unilaterally rejecting and seizing assets is exactly given by the full cash flows of the loan portfolio. Then, as a consequence, there is at least one investor who finds it profitable to reject and seize assets if the banker seeks rents during renegotiations. This causes a 'chain reaction' among investors, and the banker is disintermediated.

Further, the priority structure among investors prevents any coordination failures among the investors that lead to the disintermediation of the banker. The usual logic around bank runs—everybody rushing to the bank to get their money out before others do so—does not apply. Indeed, the senior investor seizes assets with priority. The junior investor, therefore, can never outrun the senior investor. The run equilibrium unravels.

Finally, the contract C^* specifies a small fee paid by investors seizing assets. We stress that equilibrium multiplicity arises if the banker cannot enforce such a fee. However, these equilibria are unstable: even an infinitesimal fee would unravel them.

5 Demand deposits

Here, we revisit the contract in Diamond and Rajan (2000), a combination of demandable debt issued to multiple depositors and capital issued to a single other investor. We show that this contract fails to render the loan portfolio fully liquid—in contrast to contract C^* from the previous section.[15]

Suppose the banker has issued $D > 0$ demand deposits on date 1 to a large number n of depositors, each holding a claim with a fixed face value of $\frac{D}{n}$. Depositors can demand to be repaid at every instant, withdrawing shares of the loan portfolio with a market value of $\frac{D}{n}$. The rationing mechanism is *sequential service*: those requesting to withdraw are assigned positions in the withdrawal order through a fair lottery. When a depositor gets a chance to withdraw, he can seize assets with a market value equal to the face value of the depositor's claim—as long as any assets are left. The remaining depositors receive nothing once the banker runs out of assets.

Diamond and Rajan (2001) show that deposits cannot be renegotiated down, as each investor faces strict incentives to withdraw their claim at full face value when the banker

[15] The loan portfolio is not risky in Diamond and Rajan (2001), and hence there is no such role for capital. However, as explained below, the bank may fail due to panic runs.

offers anything less. This prevents renegotiations in good states at date 2, which the banker finds desirable at date 1. However, it also prevents renegotiations in bad states at date 2, leading to runs if the banker cannot repay all depositors. Diamond and Rajan (2000) argue that capital should act as a buffer when the loan portfolio is sufficiently risky.

So, suppose the banker has issued non-demandable debt with an arbitrarily large face value. As Diamond and Rajan (2000), we interpret this claim as *capital.* Demand deposits are senior to capital. Only once depositors have been made whole can the banker repay the capital investor. The capital investor has the non-contingent right to seize shares of the assets but only up to an amount that ensures the full repayment of deposits.[16]

The main results of this section are summarized in Proposition 1, which states the following. In bad states of the economy, depositors run on the bank, knowing that not all of them can be repaid in full. In good states, there are no runs as the banker can always fully repay all those that withdraw. In intermediate states, runs may but need not occur. The loan portfolio is thus *illiquid* for any level of D since the banker cannot pledge to always repay what she collects.

Proposition 1. *The demand deposit contract generates the following equilibrium outcomes of the renegotiation game:*

1. *If $x(\theta) < D$, the equilibrium is unique and such that all depositors run on the bank, leading to disintermediation. The depositors collect a total of $\beta x(\theta)$ from the loan portfolio.*

2. *If $x(\theta) \in \left[D, \frac{D}{\beta}\right)$, there exist two equilibria. In the first equilibrium, all depositors run on the bank, leading to disintermediation, and collect a total of $\beta x(\theta)$ from the loan portfolio. In the second equilibrium, there is no run. The banker collects $x(\theta)$ from the loan portfolio and repays a total of D to the depositors and $\beta(x(\theta) - D)$ to the capital investor.*

3. *If $x(\theta) \geq \frac{D}{\beta}$, the equilibrium is unique and such that no depositor runs on the bank. The banker collects $x(\theta)$ from the loan portfolio and repays a total of D to the depositors and $\beta(x(\theta) - D)$ to the capital investor.*

The proof is in Appendix B. The intuition is as follows. First, suppose the loan portfolio's cash flows are so low that not all depositors can expect to be repaid in full. Each depositor

[16]Equivalently, we can interpret this claim as long-term debt: the long-term debt investor can seize shares of the loan portfolio if the banker defaults and only up to the amount that ensures full repayment of deposits.

then finds it optimal to run on the bank. If few other depositors withdraw, then a withdrawing depositor is made whole. If many other depositors withdraw, then not withdrawing yields a zero payoff. Withdrawing yields a strictly positive expected payoff, as this depositor may be serviced ahead of other depositors. Hence, the deposit contract admits *fundamental runs* in bad states.

Second, if the loan portfolio's cash flows are high enough, then no run occurs. Even if all other depositors withdraw, the banker still has enough shares of the loan portfolio remaining to repay an individual depositor in full, thus convincing him not to withdraw. The run equilibrium unravels. While the banker can prevent a run, she cannot renegotiate down any depositor's claim: each depositor withdraws at face value should the banker try to do so. Issuing deposits thus allows the banker to commit her skills to collect the loans on behalf of the investors in the absence of runs.

This is not true for the capital investor, who—should the banker threaten to withdraw her human capital—can seize assets as long as the depositors are repaid in full. The capital investor can thus generate $\beta(x(\theta) - D)$ by seizing assets and optimally accepting any weakly larger offer. The banker, therefore, offers to collect the loans also on behalf of the capital investor but to retain $(1 - \beta)(x(\theta) - D)$ for herself, which the investor accepts. In sum, the banker prevents runs and collects the loans in full; the depositors are repaid in full, whereas the capital investor's claim is renegotiated down.

Third, if the loan portfolio's cash flows are in an intermediate region, depositors can coordinate on running on the bank. We refer to such a scenario as *panic* runs. The logic is similar to that in Diamond and Dybvig (1983) and rests on the idea that the banker has more demandable claims outstanding than the market value of its loans. If all other depositors do not withdraw, the banker can generate sufficient resources to convince an individual depositor not to withdraw. As in the previous case, the banker collects the loans in full and repays all depositors but renegotiates with the capital investor. However, if all other depositors withdraw, the banker can generate too few resources to convince an individual depositor not to withdraw. The individual depositor then prefers withdrawing, as the nature of sequential service gives him a chance of being serviced ahead of other depositors and getting repaid in full.

What if the banker tries to eliminate panic runs by replacing old with new depositors? Recall that there is no shortage of liquidity in the market. The banker must refinance from multiple new depositors for this scheme to work. If the banker replaces withdrawing depositors with only one investor, this investor will suffer from the banker's rent-seeking as outlined in subsection 2.3. With multiple investors needed, panic-based runs do not disappear. Indeed, if all present depositors run on the bank—and potential new depositors coordinate on not replacing old depositors—the bank fails.

Furthermore, the banker cannot improve outcomes by including a withdrawal fee in the demand deposit contract. If the fee is infinitesimal, it does not prevent runs, as depositors strictly (rather than weakly) prefer running. A substantial withdrawal fee may prevent runs but allows the banker to seek rent ex-post.

Finally, note that panic runs continue to plague deposits even if agents could contract on the state θ. The first-come, first-served nature of deposits still leaves the bank vulnerable to runs even if the face value could be adjusted to the state of the economy. State-contingency does not rule out coordination failures.

Contrast to optimal contract. It is useful to contrast the demand deposit contract with the optimal contract C^*. First, depositors withdraw a variable portion of the banker's assets with a fixed value given by the face value of deposits. Under contract C^*, investors withdraw a fixed portion of the assets with a variable value.

Second, both contracts feature liquidity creation: the market value of the bank's assets is less than what bank creditors can demand to be repaid. Under the deposit contract, withdrawing depositors are serviced sequentially in a randomly determined order. Consequently, they are treated heterogeneously during a run despite being ex-ante and ex-post identical in all other dimensions. This is also true for contract C^*, which prioritizes the senior investor's claim, who is serviced first when multiple investors request to withdraw. The key difference is that contract C^* treats investors according to pre-specified rules. Therefore, investors withdraw in an 'orderly fashion.' On the other hand, runs allocate assets to depositors who manage to withdraw faster than other depositors, i.e., due to luck. Thus, depositors withdraw in a 'disorderly fashion.'

The equilibrium outcomes for the two contracts differ widely. Under contract C^*, the banker collects the loans on behalf of the investors in full and repays a total of $x(\theta)$ in every state θ. The demand deposit contract can achieve this outcome only if the face value of deposits is given by the highest possible realization of cash flows. However, issuing this maximal level of deposits necessarily leads to fundamental runs in all other states of the economy. It further leaves the bank susceptible to panic runs in the best state. This highlights a trade-off between deposits and equity: the banker can reduce the run risk by issuing fewer deposits and more capital, yet more capital inevitably leads to more ex-post rent-seeking in the absence of runs. This trade-off is not present under the optimal contract.

6 Discussion

6.1 Alternative optimal contracts

The optimal contract C^*, which did not specify promised repayments and allowed investors to seize assets even without default, is by no means unique.[17] Consider the following alternative contract C'. The banker borrows from $n \geq 2$ investors. Let x_{max} denote the highest level of cash flows that the loan portfolio could generate. Investor 1 is promised a repayment $D_1 = \beta x_{max}$. Each junior investor i is promised a repayment $D_i = \frac{(1-\beta)}{n-1} x_{max}$. Investors can seize assets in the same proportions specified by contract C^* but only if the banker defaults. Investor 1 is again senior. Seizing assets incurs a vanishing rejection fee. The results for contract C' are the same as for contract C^*.

Furthermore, the small rejection fee was paid by any investor seizing assets regardless of the banker's offer, but this is not critical either. Consider contract C'' as contract C' above but with a modified rejection fee. The fee is common across all investors and is strictly positive but arbitrarily small if $\tilde{B}_i(x)/D_i = \tilde{B}_j(x)/D_j$ for all investors $i \neq j$. Otherwise, the fee is zero. Investors only incur the rejection fee if the banker defaults by attempting to renegotiate a larger haircut for some investors relative to others. As for contract C^*, any attempt at rent-seeking induces disintermediation. Panic runs without rent-seeking cannot occur, as junior investors have no hope of 'outrunning' the senior investor and prefer not to pay the rejection fee.

6.2 Sequential service is detrimental to liquidity creation

Our analysis highlights that rationing through sequential service is detrimental to liquidity creation. First, if banks are structured to prevent rent-seeking (and hence enhance liquidity creation), Section 5 demonstrates that the deposit contract is dominated by contract C^*, which does not feature sequential service. Second, the investors in this framework are not seeking liquidity insurance since they do not have an 'urgent need to consume' as in Diamond and Dybvig (1983). Consequently, sequential service is not essential: the bank can collect

[17] An advantage of this arrangement is that default can be broadly defined to include any attempts at rent extraction, including those that cannot be described ex-ante and are thus not verifiable ex-post.

all withdrawal requests and then service withdrawing investors in some contractually pre-specified order. Even if some investors urgently need to consume, there is no aggregate liquidity shortage.[18] Given the strength of the investors' claims under contract C^*, financial markets (i.e., storing investors) should be willing to purchase claims on the bank if other investors seek to sell them.

An important reason for sequential service is that it incentivizes investors to monitor the bank (Calomiris and Kahn, 1991). Specifically, monitoring is privately costly, and not all investors are equally good at it. Sequential service incentivizes those good at monitoring to withdraw at the first sign of trouble and be repaid in full (their reward for being vigilant), which would then alert the non-monitoring investors that they should also withdraw. The banker is deterred from acting against the investors' preferences. In contrast, suppose instead of sequentially, all bank creditors are repaid pro-rata in states when the bank fails. This arrangement leads to a free rider problem since monitoring gains are equally shared among monitors and non-monitors.

However, first-come, first-serve may also trigger a preemptive run by uninformed investors in the spirit of Chen (1999). Specifically, the monitoring investors would withdraw earlier than the rest in bad states when the bank cannot repay in full. To address this informational disadvantage, the non-monitoring investors may respond to other sources of information—such as a failure of another bank perceived to be similar—before the value of their bank's assets is revealed. The monitoring investors recognize that and also run on the bank, even if their signal is good. This situation can lead to the failure of healthy banks.

Notice that the monitoring incentives remain intact under our contract C^*, which does not rely on sequential service but would rather designate those investors who monitor as senior (see Section 7.2). The key observation is that overcoming the free-rider problem associated with monitoring the bank requires prioritizing some investors' claims over others, which naturally emerges under our contract. Specifically, the senior investors will be rewarded for their due diligence without the risk of preemptive runs. The reason is that monitors will run on the bank after having a good signal only when they expect the non-monitors to run—as with sequential service. However, the non-monitors can never outrun the monitors since the latter holds the senior claims. The preemptive run equilibrium in the spirit of Chen (1999) unravels.

[18] As Diamond and Rajan (2000) put it, "the bank is a source of liquidity both for the depositor and the entrepreneur. When some (or even all) initial depositors want their money back in the ordinary course of business (in contrast to a run), the bank does not need to liquidate the entrepreneur. It simply borrows from new depositors who, given the prospective strength of their claim, will willingly refinance." (p. 2432)

6.3 Safe claims and payment services

The optimal contract C^* in Section 3 implies that the investors hold risky claims on the bank. What if (some) investors want (at least partially) safe claims, e.g., for payment services in the spirit of Gorton and Pennacchi (1990) and Dang et al. (2017)?

First, while the losses imposed by C^* may disrupt payments, it is certainly true that the bank runs induced by the deposit contract bring about a much greater disruption. Second, the banker can engineer a safe claim by tranching the loan portfolio. The senior tranch generates a cash flow of x_{min} in every state, where $x_{min} > 0$ is the lowest possible payoff of the loan portfolio. The junior tranch generates a cash flow of $x(\theta) - x_{min}$ in state θ. One can think of the banker as creating two subsidiaries: one fully safe (the senior tranch) and one bearing all risks (the junior tranch). Each of these subsidiaries (or tranches) is financed with contract C^*. Consequently, all claims on the senior tranch have a fixed value since the banker cannot seek rents, and the senior tranch is constructed to feature no real asset risk. Investors seeking a safe claim invest in the senior tranch, while others invest in the junior tranch.

What if investors want claims that are fully safe *and* obey sequential service, e.g. for reasons as in Wallace (1988)? Contract C^* then cannot be used to finance the senior tranch. In that case, the banker can issue first-come, first-served deposits with a total face value of βx_{min} and finance the remainder with contract C^* (adjusted to ensure that depositors take priority). Applying our Theorem and Proposition 1, runs then do not occur in any state and any equilibrium. The deposits are fully safe and can serve as payment vehicles, and the banker cannot seek rents from any investor. This observation further serves to highlight that contract C^* not only improves upon the deposits but also upon capital (or long-term debt) as a safety buffer.

6.4 Implementation and government distortions

Our contract C^* is a mild modification of the existing deposit contract along two dimensions. First, investors withdraw a fixed portion of shares rather than a variable portion of shares with a fixed value. Second, investors are not serviced sequentially. Instead, the bank collects all withdrawal requests and services some claims with priority. The contract should, therefore, be readily implementable in practice.

Further, Flannery (1994) finds that agency costs are particularly important for banks.

Correia et al. (2024) show that conflicts of interest and fraud were prominent in bank failures. Even if banks' primary goal is issuing safe assets (e.g., for payments as in Gorton and Pennacchi (1990)) or providing liquidity insurance as in Diamond and Dybvig (1983), Subsection 6.3 argued that banks should combine deposits with the contract C^* as long as agency costs are a concern. Hence, the contract C^* should be desirable under very general circumstances.[19]

So why is the contract C^* not being widely used? One reason is government guarantees. Specifically, while fragility is detrimental to liquidity creation, government guarantees could make it optimal for banks to finance with a fragile liability structure. Specifically, it is well-known that deposit insurance and government bailouts distort banks' incentives (see, e.g., Farhi and Tirole (2012); Keister (2016); Dávila and Walther (2020); Philippon and Wang (2023)). This remains true in our case as well. Indeed, one of the main properties of C^* is that bank creditors take losses in bad states of the economy. This is reminiscent of a bail-in. However, banks may choose a non-bail-inable liability structure or may be deterred from bailing in their creditors if they anticipate being eventually bailed out (see, e.g., Keister and Mitkov (2023)). Similarly, deposit insurance may make it privately optimal for banks to issue rigid claims such as demand deposits in excessive amounts (see, e.g., Kareken (1985); Georgiadis-Harris and Guennewig (2024)). In both cases, banks maximize the expected government subsidy. In our context, this implies that they could use C^* but choose not to. Fragility may, therefore, be an undesirable side effect of government interventions and not an undesired but necessary disadvantage of liquidity creation.

7 Extensions

7.1 State-contingent loan collection skills

We assumed in the baseline model that the bank's superior loan collection skills captured by β were fixed. At the same time, the optimal contract in Section 3 depends on β through the portion of the banker's assets each junior investor can seize. Here, we show that our results apply even if β varies with the state of the economy. So, assume β is a random variable

[19] Another piece of evidence favoring the agency cost view is that, historically, suspensions have not been an option for individual banks. Instead, they have been a regulatory response to system-wide panics. This is consistent with our setup, where bankers would misuse an option to suspend payments, and inconsistent with Diamond and Dybvig (1983), where an option to suspend should be available to banks (Wallace, 1988; Andolfatto et al., 2017).

governed by the CDF $G(.)$ with support in $(0, 1)$ where G is common knowledge.[20]

First, assume the realized value of β is observable and verifiable. Hence, the portion of assets each junior investor can seize can be contingent on the realization. Recall that β is also the market value of a loan unit at date 1. Therefore, it is not unreasonable to believe that the realization of β is verifiable. In this case, our results from Section 3 apply verbatim.[21]

The more interesting case is when β is observable but not verifiable, implying that the initial contract cannot be contingent on the realized value of β. To simplify, we restrict the domain of β to the interval $(\frac{1}{2}, 1)$ and focus on two investors: senior and junior. At the end of this section, we explain how to extend the logic to the full interval, which requires a larger number of junior investors.

As before, the senior investor can seize a portion of $s_1 = 1$ of the banker's assets. What is different is that the portion that the junior investor can seize s_2 is now left unspecified—the contract only provides *rules* for how this value will be determined later. Specifically, before the renegotiation game begins on date 2, the junior investor observes the realized value of β and must unilaterally select a value of $\tilde{\beta}$ from the interval $(\frac{1}{2}, 1)$. His choice is revealed to all parties. The portion of assets the junior investor can seize is then $s_2 = (1 - \tilde{\beta})/\tilde{\beta}$ and the renegotiation game begins.

The junior investor recognizes that the selected value of $\tilde{\beta}$ affects the behavior of the senior investor and the banker in the resulting renegotiation game. Suppose the junior investor selects the correct value $\tilde{\beta} = \beta$. We know from Section 3 that the junior investor's payoff in the corresponding renegotiation game is unique and given by $(1 - \beta)x$. We will show that any other choice of $\tilde{\beta}$ leads to a strictly lower payoff for the junior investor.

Specifically, recall that the junior investor gets a payoff of zero whenever the senior investor rejects. Moreover, the senior investor will not accept unless the junior investor accepts (see Lemma 1). Suppose the junior investor selects $\tilde{\beta} < \beta$. The senior investor will not accept unless offered at least βx since he can always reject, seize all assets, and get βx. Hence, the junior investor can be offered at most $(1 - \beta)x$ when the senior investor accepts. But then the junior investor's rejection payoff is $(1 - \tilde{\beta})x$, which is greater than $(1 - \beta)x$.

[20]Assuming that investors can collect a (random) proportion $\beta \in (0, 1)$ of what the banker can collect is without loss of generality. Indeed, suppose that investors can collect some $y(\theta) < x(\theta)$. Defining $\beta(\theta) \equiv y(\theta)/x(\theta)$ then recovers the proportionality. Note that our result holds for any realization of β.

[21]At the same time, contracting on market signals is challenging (see Section 5 in Mitkov (2024)). Market-based contracts can be open to manipulation (Diamond, 1993), lead to equilibrium multiplicity (Sundaresan and Wang, 2015; Pennacchi and Tchistyi, 2019), or incentivize the parties to act in ways that diminish their informativeness (Bond et al., 2010). Perhaps for those reasons, Chava and Roberts (2008) finds that debt covenants are seldom based on market signals.

Therefore, the banker cannot get both investors to accept, implying that the junior investor achieves a zero payoff. Second, suppose the junior investor selects $\tilde{\beta} > \beta$. Then, the junior investor's payoff is less than $(1-\beta)x$ in the corresponding renegotiation game. Indeed, by rejecting the junior investor, it achieves payoff bound from above by $(1-\tilde{\beta})x < (1-\beta)x$. Hence, the banker offers βx to the senior investor less than $(1-\beta)x$ to the junior investor, which both accept.

Thus, if $\beta < \tilde{\beta}$, the junior investor is penalized by the senior investor; if $\beta > \tilde{\beta}$, he is penalized by the banker. Hence, the junior investor sets $\tilde{\beta}$ to the realized value of β. A natural interpretation is that the junior investor voluntarily *bails in*. Also, the junior investor must select $\tilde{\beta}$—not the banker or senior investor. The banker will select $\tilde{\beta} > \beta$ since this allows her to repay less than $(1-\beta)x$ to the junior investor and thus earn rents ex-post. Similarly, the senior investor has no incentive to select the correct value of β since his equilibrium payoff is always at least βx, regardless of what he selects. He may, therefore, be tempted to seek rents by threatening to set the junior investor's share to something strictly less than $(1-\beta)$.

Finally, suppose there are multiple junior investors. This is necessary if the lower bound of the support for β is less than $\frac{1}{2}$. For example, suppose $\beta \in (\beta_1, \beta_2)$ where $0 < \beta_1 < \beta_2 < 1$. Denote by n the smallest integer such that $\beta_1 \geq \frac{1}{n}$. The optimal contract then requires at least $n-1$ junior investors. The optimal contract has one of those junior investors choose $\tilde{\beta}$ for all junior investors and is otherwise given by C^*. The reasoning is perfectly analogous to the reasoning in the case of one junior investor. Selecting a level of $\tilde{\beta}$, which is too high, allows the banker to seek rents; selecting a level that is too low induces the senior investor to seize all assets. Setting $\tilde{\beta}$ to the realization β prevents both.

7.2 Monitoring

We mentioned in Section 6.2 that one desirable property of demand deposits subject to first-come, first-served is that it incentivizes the depositors to monitor the bank. Investors do not need to monitor the banker in our baseline model since the state θ is perfectly observable. We now introduce a role for monitoring and show that the contract C^* preserves investors' monitoring incentives. To illustrate, we revisit the two-investor example from Section 4 but assume the state θ is unobservable to the investors unless they monitor the bank. Monitoring reveals the state perfectly but is privately costly. For simplicity, we assume that monitoring is contractible. Below, we show that the contract C^* remains optimal with one modification:

the senior investor is designated as the monitor.

As before, the banker cannot extract rents. Indeed, if the state is θ, the senior investor (who monitors and thus knows the state) must be offered a payment of at least $\beta x(\theta)$. After observing the offer to the senior investor, the junior infers the state is θ and would not accept unless offered a payment of at least $(1-\beta)x(\theta)$. In other words, the offer to the informed senior investor conveys information to the uninformed junior investor. Insisting on the feasibility of the offered repayments, the only way for the banker not to get disintermediated is to offer $\beta x(\theta)$ to the senior investor and $(1-\beta)x(\theta)$ to the junior investor for a total repayment of $x(\theta)$. The banker can thus raise $\mathbb{E}[x(\theta)] - \mu$ by using the loan portfolio as collateral, where μ is the senior investor's monitoring cost.

Finally, recall that liquidation is an off-equilibrium threat in the optimal contract. One can modify the model to generate liquidation as part of the optimal contract and show that C^* remains optimal. Assume, for simplicity, that there are only two states: good and bad. In the bad state, the investors can collect more by liquidating the bank rather than continuing. For example, this situation arises in Calomiris and Kahn (1991) because the banker absconds with a portion of the revenues when left in charge when the state is bad. As before, we use the contract C^* with the senior investor appointed as the monitor. The bank will then be liquidated by the senior investor in the bad state.

8 Full framework

This section outlines the full framework of financial intermediation, which captures the essential ingredients of Diamond and Rajan (2000, 2001). The full framework incorporates an entrepreneur to whom the banker issues a loan to finance productive investment. Looking ahead, it is an outcome of the model that the investors collect the loan directly from the entrepreneur should the banker be disintermediated. It, therefore, remains optimal for the banker to select contract C^*, as this allows her to avoid disintermediation and pledge the full loan cash flow in every equilibrium and every state to the investors.

Intuition. We start with high-level intuition. Renegotiation comprises two stages. The first determines who owns the project's assets—possibly through a run by depositors as in Diamond and Rajan (2000, 2001), or the investors' decisions to seize assets under contract C^*. The second determines who gets to operate these assets. The entrepreneur remains the first-best user of the project's assets. This is in contrast to the banker, who is only the second-best user and, therefore, does not create value. Instead, she only facilitates transfers from one agent to another. Optimally, the entrepreneur cuts out the banker by offering to repay to the owners of the asset whatever the banker would offer if she were rehired to deal with the entrepreneur. The banker thus prefers to avoid disintermediation in the first place. This is in contrast to the entrepreneur, who, as the first-best user of the assets, always ends up operating them. Consequently, the threat of disintermediation disciplines the banker but not the entrepreneur.

The remainder of this section fills in the details.

Framework. An entrepreneur has access to a project that requires a unit investment but lacks the necessary funds. Investment is contractible. If the entrepreneur works on the project on date 2, it generates some cash flows $c(\theta) > 0$ in state θ; if the entrepreneur does not work, it generates zero cash flows. A relationship lender (the banker) finances the entrepreneur and, in the process, develops specific skills in redeploying the projects' assets to their second-best use. This second-best use generates returns $x(\theta) < c(\theta)$ at date 2. Finally, there are many other investors, each of whom can liquidate the project's assets to generate returns $\beta x(\theta)$, with $\beta \in (0, 1)$. Neither entrepreneur nor banker can commit their future human capital (but both can commit it in a given period). The state θ is observable but not verifiable. Cash flows are both observable and verifiable. The project's assets and, thus, the loan to the entrepreneur are perfectly divisible.

The banker may again experience a liquidity shock on date 1. If so, she wants to borrow as much as possible from the (deep-pocketed) investors using the loan as collateral. Alternatively, the banker borrows from the investors on date 0 since she lacks sufficient funds to finance the entrepreneur. The banker's precise motivation to borrow is unimportant, and the model's insights are unchanged. The state θ is realized at the beginning of date 2, at which point the renegotiation game begins. Figure 3 summarizes the timeline of events in the full framework.

Note that illiquidity is socially costly. The investment is socially desirable if $\mathbb{E}[c(\theta)] > 1$. However, if the loan to the entrepreneur is illiquid and the banker experiences the liquidity shock with a sufficiently high probability, then the entrepreneur cannot obtain financing if $\beta\mathbb{E}[x(\theta)] < 1$.

Renegotiation. After the state θ is realized, the banker, the investors, and the entrepreneur play the enhanced renegotiation game. As before, the banker moves first by offering a repay-

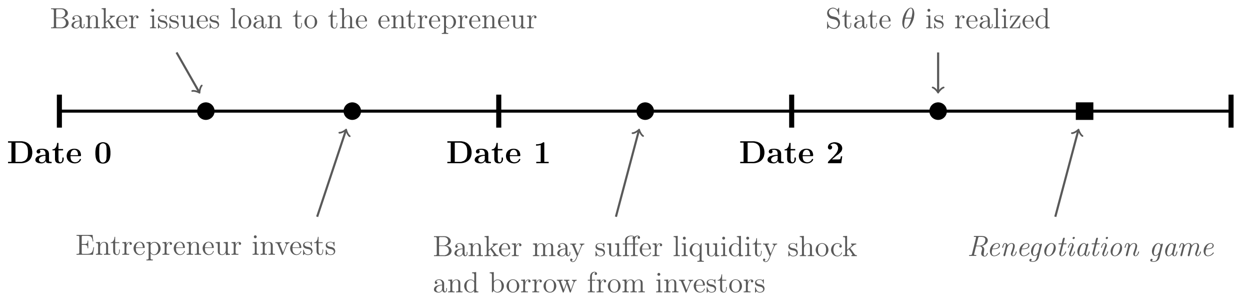

Figure 3: Timeline of events in the full framework.

ment to the investors. If the investors accept this offer, the banker renegotiates the loan terms with the entrepreneur, collects the loan, and repays the investors. The game ends. Investors can also reject the offer and seize assets, i.e. (parts of) the loan to the entrepreneur. They then negotiate the terms of the loan directly with the entrepreneur. Should they come to an agreement, the investors directly collect the loan, and the game ends. Should they fail to reach an agreement, the investors can rehire the banker to negotiate with the entrepreneur and collect the loan on their behalf. Should they decide not to do so, they liquidate the project's assets. Throughout the game, the banker and entrepreneur make take-it-or-leave-it offers to the investors, and the entrepreneur makes take-it-or-leave-it offers to the banker. (Again, this is without loss of generality.) Figure 4 illustrates the enhanced renegotiation game.[22]

Mapping to our simpler framework. Once investors have seized assets, the outcome of the renegotiation game is the following. Since the investors can only generate $\beta x(\theta)$ when liquidating a unit of the project's assets, the banker offers $\beta x(\theta)$ to collect the loan at the final node. Investors cannot do better and accept. Anticipating that investors accept an offer of $\beta x(\theta)$, the entrepreneur offers to repay $\beta x(\theta)$. The investors cannot do better by rejecting this offer and accept. From here, it follows that—as in our main analysis—investors only collect a share β of the potential cash flows after the banker has been disintermediated. Furthermore, the banker is not rehired after disintermediation, implying a zero payoff. Finally, when the banker collects the loan, the entrepreneur offers $x(\theta)$. Since the banker cannot do better by

[22] We have adopted a specific sequence of events in Figure 4, but this is not crucial. What drives the results is that the entrepreneur generates value ex-post, whereas the banker only facilitates the transfer of resources from one agent to another.

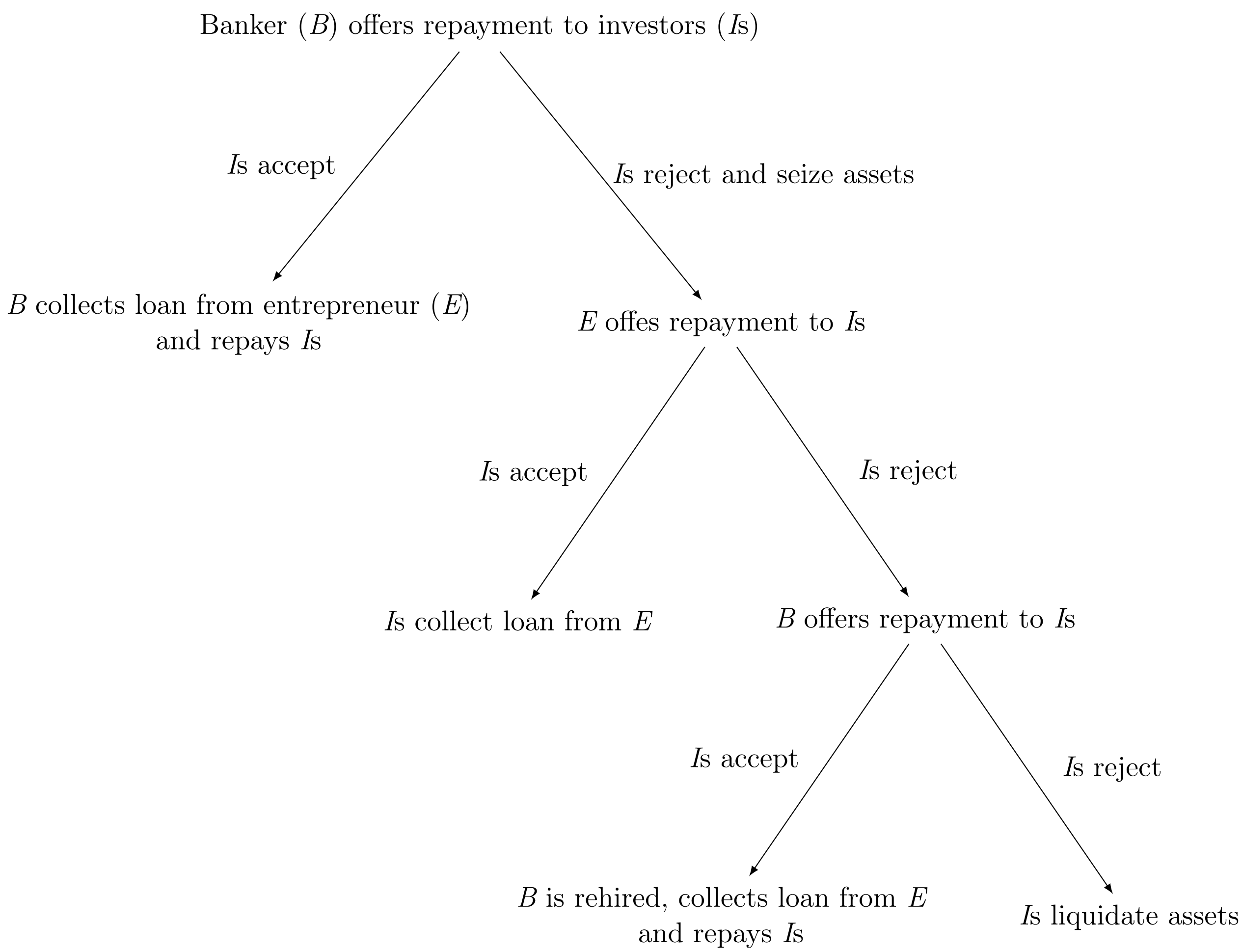

Figure 4: Sequence of events within the enhanced renegotiation game.

seizing the project's assets and putting them to their second-best use, she accepts. It then follows that the full framework maps directly into our simplified framework of Section 2 and that contract C^* remains optimal.

9 Conclusion

Contrary to previous findings in the literature (Diamond and Rajan, 2000, 2001), we show that banks should not adopt a run-prone financing structure. Instead, financial fragility is *detrimental* to liquidity creation. We characterized an optimal contract that rests on two simple notions: bank creditors can seize pre-specified portions of the bank's assets, and some creditors are treated preferentially in doing so. Our contract prevents rent-seeking, ensures the parties renegotiate only when necessary to absorb real asset losses, and averts coordination failures where each investor rushes to seize assets expecting the same from the others. We also argued that our contract is easily implementable. However, banks may be deterred from adopting it due to government guarantees incentivizing excessive reliance on demand deposits.

References

Allen, F. and D. Gale (2000): "Financial contagion," *Journal of political economy*, 108, 1–33.

Andolfatto, D., E. Nosal, and B. Sultanum (2017): "Preventing Bank Runs," *Theoretical Economics*, 12, 1003–1028.

Berglöf, E. and E.-L. Von Thadden (1994): "Short-term versus long-term interests: Capital structure with multiple investors," *The quarterly journal of economics*, 109, 1055–1084.

Bhattacharya, S. and D. Gale (1987): "Preference Shocks, Liquidity, and Central Bank Policy," in *New Approaches to Monetary Economics (edited by William A. Barnett and Kenneth J. Singleton)*, New York: Cambridge University Press, 69–88.

Bolton, P. and D. S. Scharfstein (1996): "Optimal debt structure and the number of creditors," *Journal of political economy*, 104, 1–25.

Bond, P., I. Goldstein, and E. S. Prescott (2010): "Market-based corrective actions," *The Review of Financial Studies*, 23, 781–820.

Bris, A. and I. Welch (2005): "The optimal concentration of creditors," *The Journal of Finance*, 60, 2193–2212.

Calomiris, C. W., C. Kahn, and S. Krasa (1991): "Optimal contingent bank liquidation under moral hazard," *BEBR faculty working paper; no. 91-0159.*

Calomiris, C. W. and C. M. Kahn (1991): "The role of demandable debt in structuring optimal banking arrangements," *The American Economic Review*, 497–513.

Chava, S. and M. R. Roberts (2008): "How does financing impact investment? The role of debt covenants," *The journal of finance*, 63, 2085–2121.

Chen, Y. (1999): "Banking panics: The role of the first-come, first-served rule and information externalities," *Journal of political economy*, 107, 946–968.

Correia, S. A., S. Luck, and E. Verner (2024): "Failing Banks," *NBER Working Paper w32907*.

Dang, T. V., G. Gorton, B. Holmström, and G. Ordonez (2017): "Banks as secret keepers," *American Economic Review*, 107, 1005–1029.

Dávila, E. and A. Walther (2020): "Does size matter? Bailouts with large and small banks," *Journal of Financial Economics*, 136, 1–22.

Diamond, D. and R. Rajan (2000): "A Theory of Bank Capital," *Journal of Finance*, 55, 2431–2465.

——— (2001): "Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking," *Journal of Political Economy*, 109, 287–327.

Diamond, D. W. (1984): "Financial intermediation and delegated monitoring," *The review of economic studies*, 51, 393–414.

——— (1993): "Seniority and maturity of debt contracts," 33, 341–368.

——— (2004): "Presidential address, committing to commit: short-term debt when enforcement is costly," *The journal of Finance*, 59, 1447–1479.

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

Diamond, D. W. and R. G. Rajan (2005): "Liquidity shortages and banking crises," *The Journal of finance*, 60, 615–647.

Farhi, E. and J. Tirole (2012): "Collective moral hazard, maturity mismatch, and systemic bailouts," *American Economic Review*, 102, 60–93.

Flannery, M. J. (1994): "Debt maturity and the deadweight cost of leverage: Optimally financing banking firms," *The American economic review*, 84, 320–331.

Fudenberg, D. and J. Tirole (1991): *Game Theory*, MIT press.

Georgiadis-Harris, A. and M. Guennewig (2024): "Bank Resolution, Deposit Insurance, and Fragility," *CRC TR 224 Discussion Paper No. 477/2023*.

Gorton, G. and G. Pennacchi (1990): "Financial intermediaries and liquidity creation," *The Journal of Finance*, 45, 49–71.

Hart, O. and J. Moore (1994): "A theory of debt based on the inalienability of human capital," *The Quarterly Journal of Economics*, 109, 841–879.

Holmström, B. and J. Tirole (1998): "Private and public supply of liquidity," *Journal of political Economy*, 106, 1–40.

Ivashina, V., B. Iverson, and D. C. Smith (2016): "The ownership and trading of debt claims in Chapter 11 restructurings," *Journal of Financial Economics*, 119, 316–335.

Jackson, T. H. (1986): *The logic and limits of bankruptcy law*, Beard Books.

Kareken, J. H. (1985): "Ensuring financial stability," in *Federal Reserve Bank of San Francisco Proceedings*, June, 53–86.

Keister, T. (2016): "Bailouts and financial fragility," *The Review of Economic Studies*, 83, 704–736.

Keister, T. and Y. Mitkov (2023): "Allocating losses: Bail-ins, bailouts and bank regulation," *Journal of Economic Theory*, 210, 105672.

Lipsey, R. G. and K. Lancaster (1956): "The general theory of second best," *The review of economic studies*, 24, 11–32.

Mitkov, Y. (2024): "A Theory of Debt Maturity and Innovation," *Journal of Economic Theory*, 218, 105828.

Pennacchi, G. (2012): "Narrow banking," *Annu. Rev. Financ. Econ.*, 4, 141–159.

Pennacchi, G. and A. Tchistyi (2019): "On equilibrium when contingent capital has a market trigger: A correction to Sundaresan and Wang Journal of Finance (2015)," *The Journal of Finance*, 74, 1559–1576.

Philippon, T. and O. Wang (2023): “Let the worst one fail: a credible solution to the too-big-to-fail conundrum,” *The Quarterly Journal of Economics*, 138, 1233–1271.

Sundaresan, S. and Z. Wang (2015): “On the design of contingent capital with a market trigger,” *The Journal of Finance*, 70, 881–920.

Von Thadden, E.-L., E. Berglöf, and G. Roland (2010): “The design of corporate debt structure and bankruptcy,” *The Review of Financial Studies*, 23, 2648–2679.

Wallace, N. (1988): “Another attempt to explain an llliquid banking system: With sequential service taken seriously,” *Federal Reserve Bank Minneapolis Quarterly Review*, 3–15.

Appendix

A Proof of Lemma 2

Fix the rejection fee κ and the banker's offer $(\tilde{B}_i(x))_{i=1}^n$. Let $p_i(\kappa) \in [0,1]$ denote the corresponding equilibrium probability that investor i accepts the banker's offer. The banker is disintermediated whenever the senior investor rejects, which happens with probability $1 - p_1(\kappa)$. The senior investor can always get a payoff of $B_1^*(x) - \kappa$ by rejecting since his seniority ensures that he always seizes all assets. So, suppose the banker offers at least $B_1^*(x) - \kappa$ to the senior investor but offers less than $B_i^*(x) - \kappa$ to at least one junior investor.

$$\tilde{B}_1(x) \geq B_1^*(x) - \kappa \qquad \text{and} \qquad \tilde{B}_i(x) < B_i^*(x) - \kappa \qquad \text{for some } i \geq 2.$$

We will show that as the rejection fee κ vanishes, the probability of the senior investor accepting goes to zero: $\lim_{\kappa \to 0} p_1(\kappa) = 0$.

Denote by $\alpha_{-i} \in \{0, ..., n-2\}$ the number of junior investors other than i accepting the banker's offer. With a slight abuse of notation, we let α_{-i} also denote the random variable induced by the players' strategies. Thus, $P\{\alpha_{-i} = m\}$ is the probability that m of the other junior investors accept the banker's offer. Consider some junior investor i who has received an offer $\tilde{B}_i(x) < B_i^*(x) - \kappa$. We refer to this investor as the ‘junior investor’ for the remainder of the proof. The junior investor's expected payoff from accepting the banker's offer is then given by

$$\mathcal{V}_i^A(\kappa) = p_1(\kappa)\left[P\{\alpha_{-i} = n-2\}\tilde{B}_i(x) + P\{\alpha_{-i} < n-2\}\mathcal{U}_i^A\right] + (1 - p(\kappa)) \cdot 0,$$

where $\tilde{B}_i(x)$ is the junior investor's payoff from accepting when all other investors accept, and $\mathcal{U}_i^A$ is his expected payoff from accepting when the senior investor accepts, but not all junior investor accept. We can express $\mathcal{U}_i^A$ as

$$\mathcal{U}_i^A = \sum_{m=0}^{n-3} \tfrac{P\{\alpha_{-i}=m\}}{P\{\alpha_{-i}<n-2\}}\mathcal{U}_i^A(m),$$

where $\mathcal{U}_i^A(m)$ is the junior investor's expected payoff from accepting when the senior investor and $m \in \{0, ..., n-3\}$ of the other junior investors accept. His expected payoff from rejecting the banker's offer is given by

$$\mathcal{V}_i^R(\kappa) = p_1(\kappa)\left[P\{\alpha_{-i} = n-2\}(B_i^*(x) - \kappa) + P\{\alpha_{-i} < n-2\}(\mathcal{U}_i^R - \kappa)\right] + (1 - p_1) \cdot (-\kappa),$$

where $B_i^*(x)$ is the junior investor's gross payoff from rejecting (excluding the fee) when all other investors accept, and $\mathcal{U}_i^R$ is his expected gross payoff from rejecting when the senior investor accepts, but not all junior investor accept. We can express $\mathcal{U}_i^R$ as

$$\mathcal{U}_i^R = \sum_{m=0}^{n-3} \tfrac{P\{\alpha_{-i}=m\}}{P\{\alpha_{-i}<n-2\}}\mathcal{U}_i^R(m),$$

where $\mathcal{U}_i^R(m)$ is the junior investor's gross expected payoff from rejecting when the senior investor and $m \in \{0, ..., n-3\}$ of the other junior investors accept.

We now show that the following relationship holds:

$$\mathcal{U}_i^A(m) \leq \mathcal{U}_i^R(m) \qquad \text{for all } k \in \{0, ..., n-3\}.$$

To this end, suppose the senior as well as m of the junior investors accept whereas $n-m-1$ of the junior investors reject and seize assets. The banker has a share $S(m)$ of the assets left, where

$$S(m) \equiv \max\left\{1 - (n-m-1)\tfrac{1-\beta}{\beta(n-1)}, 0\right\}.$$

The maximum she can pay to the investors who accepted her offer is $S(m)x$. Recall that a junior investor gets a zero repayment whenever the banker has insufficient funds to repay the senior investor.

If $S(m) \geq \beta$, then the banker has sufficient resources to repay $B_1^*(x) = \beta x$ to the senior investor. In that case, the junior investor i's payoff from accepting is at most $\mathcal{U}_i^A(m) = \tilde{B}_i(x)$. Recall that $\beta > \frac{(1-\beta)}{\beta(n-1)}$ implies that the junior investor can seize his pre-specified portion of the bank's assets if $S(m) \geq \beta$. Thus, his gross payoff from rejecting (excluding the fee) is $\mathcal{U}_i^R(m) = B_i^*(x)$. Since $\tilde{B}_i(x) < B_i^*(x)$, we have $\mathcal{U}_i^A(m) < \mathcal{U}_i^R(m)$.

The senior investor cannot be repaid $B_1^*(x) = \beta x$ if $S(m) < \beta$. In that case, the junior investor's payoff from accepting is zero. His gross payoff from rejecting (excluding the fee) is bound from below by zero. Thus, we find that $\mathcal{U}_i^A(m) \leq \mathcal{U}_i^R(m)$ for all $m \in \{0, ..., n-3\}$ and hence $\mathcal{U}_i^A \leq \mathcal{U}_i^R$.

We complete the proof as follows. If the senior investor accepts with some positive probability $p_1(\kappa) \in (0, 1]$ in equilibrium, then the junior investor must weakly prefer to accept as well: $\mathcal{V}_i^A(\kappa) \geq \mathcal{V}_i^R(\kappa)$. Otherwise, he should reject the banker's offer with probability one. But then, by Lemma 1, the senior investor also rejects with probability one. Since $\tilde{B}_i(x) < B_i^*(x)$ and $\mathcal{U}_i^A \leq \mathcal{U}_i^R$ for the junior investor, we have $\lim_{\kappa \to 0} \mathcal{V}_i^A(\kappa) < \lim_{\kappa \to 0} \mathcal{V}_i^R(\kappa)$ for all $p_1(\kappa) > 0$. Hence, it must be that $\lim_{\kappa \to 0} p_1(\kappa) = 0$. The claim follows.

B Proof of Proposition 1

Suppose the state θ is such that the loan portfolio's cash flows should the banker deploy her skills is x. Recall that the banker initially promised to repay D to the depositors. After the state is realized, renegotiation opens, the banker promises to repay $\tilde{D} \leq x$ to the depositors (where $\tilde{D}$ need not equal D). If they accept portion $s_D = \frac{\tilde{D}}{x}$ of the assets are put to the side (i.e., ringed fenced) to serve as collateral, ensuring that the depositors will be repaid $\tilde{D}$. Renegotiation then opens with the capital investor, who can seize the remaining portion $1 - s_D$ of the assets and collect $\beta x(1 - s_D)$. Thus, conditional on the depositors accepting the banker's offer, she only repays $\beta x(1 - s_D) = \beta(x - \tilde{D})$ to the capital investor. As we shall show below, if the depositors reject the banker's offer, they will seize all assets, and the payoff to both the banker and the capital investor will be driven to zero. Thus, the banker maximizes her ex-post payoff by choosing the lowest possible value of $\tilde{D}$ such that the depositors accept.

Case 1. $x < D$. The unique equilibrium is for all depositors to reject the banker's offer. Suppose all depositors accept the banker's offer and do not run. Since the banker can offer at most $\tilde{D} = x < D$, each depositor can be paid at most $\frac{\tilde{D}}{n} < \frac{D}{n}$. Each depositor has a profitable deviation. That is, since $\frac{D}{\beta xn} \leq 1$ for all n large enough, a depositor who rejects while all others accept will be made whole (i.e., he seizes $\frac{D}{\beta xn}$ units of the asset getting $\beta x \frac{D}{\beta xn} = \frac{D}{n}$). In fact, one can establish something much stronger: all depositors have a strictly dominant strategy to reject any banker's offer (see **Case 2** below). The banker cannot avert a run, and her payoff and that of the capital investor will be driven to zero. The overall payoff to the depositors is βx. Finally, what if not all depositors are promised an equal share of the cash flow x? In that case, fix a depositor who gets less than $\frac{D}{n}$. At least one such depositor exists since they cannot all be paid at least $\frac{D}{n}$. The above argument then applies.

Case 2. $D \leq x < \frac{D}{\beta}$. Suppose first the banker does not attempt to renegotiate $\tilde{d} = \frac{D}{n}$. Since $x \geq D$, she can repay $\frac{D}{n}$ to each depositor as long as they accept. That is, if all other depositors accept, each depositor will be indifferent between accepting and rejecting (he gets $\frac{D}{n}$ either way), and hence accepts. The depositors are paid D, the capital investor is paid $\beta(x - D)$, and the banker gets the rest $(1 - \beta)(x - D)$. Even if the banker does not attempt renegotiation, the depositors might coordinate on a self-fulfilling run in the spirit of Diamond and Dybvig (1983). Suppose all other depositors reject the banker's offer. The units of the asset remaining with the banker are $(1 - \frac{n-1}{n} \frac{D}{\beta x})^+$ which, for n large enough, will be zero. Thus, an investor who accepts gets zero; by rejecting, he can be made whole if lucky to be early on the line. We have constructed an equilibrium in which each depositor rejects not because the banker tries to repay less but because the other depositors are expected to reject as well. This outcome disintermediates the banker and leaves the capital investor with a zero payoff.

What if the banker attempts to renegotiate by offering to repay less than $\frac{D}{n}$ per depositor, say $\tilde{d} < \frac{D}{n}$? In that case, we will show that the equilibrium is unique and that all depositors reject the banker's offer in strictly dominant strategies. Assume exactly k depositors reject the banker's offer and seize assets, leaving the banker with $S(k) = \left(1 - k\frac{D}{\beta xn}\right)^+$ units of the asset. For n large enough, each depositor's expected payoff from rejecting is strictly greater for any k than his expected payoff from accepting. To see why, define k^* as the smallest integer with the following property: whatever assets remain after k^* depositors have seized assets is insufficient to fully repay the next depositor who gets a chance to seize assets. That is,

$$S(k^*) < \frac{D}{\beta xn} \leq S(k^* + 1) \qquad \Rightarrow \qquad k^* \equiv \left| \frac{\beta xn}{D} - 1 \right|$$

where $\left\lfloor \frac{\beta x n}{D} - 1 \right\rfloor$ denotes the smallest integer less than or equal to $\frac{\beta x n}{D} - 1$. Suppose exactly $k^* - 1$ depositors reject the banker's offer. If one more depositor rejects, his payoff will be at least $\beta x S(k^*)$ whereas if he accepts, his payoff will be $\frac{xS(k^*)}{n-k^*}$. The only thing that may deter a depositor from rejecting in this situation is $\beta x S(k^*) < \frac{xS(k^*)}{n-k^*}$ which is possible due to the banker's superior loan collection abilities. But notice that $\beta x S(k^*) > \frac{xS(k^*)}{n-k^*}$ whenever

$$n - \left\lfloor \frac{\beta x n}{D} - 1 \right\rfloor > \frac{1}{\beta}.$$

Since $\beta x < D$, the above condition is satisfied for all n large enough. Now, consider any depositor i given that exactly $k \in \{0, ..., n-1\}$ of depositors other than i reject the banker's offer. If he rejects and $k < k^* - 1$, he will be made whole. If he rejects and $k = k^* - 1$, he gets at least $\beta x S(k^*)$ and may also be made whole. Finally, if he rejects and $k > k^* + 1$, he may be made whole, get the smaller amount of $\beta x S(k^*)$, or get zero. On the other hand, if he accepts and $k < k^* - 1$, he gets at most $\tilde{d}$, which is still less than $\frac{D}{n}$ (we have assumed the banker attempts to renegotiate down the depositors). If he accepts and $k = k^* - 1$ he gets $\frac{xS(k^*)}{n-k^*}$ which is less than $\beta x S(k^*)$. Finally, if he accepts and $k > k^* - 1$, he gets zero since the banker will be left with no assets. This establishes that each depositor has a strictly dominant strategy to reject any banker's offer $\tilde{d}$ less than $\frac{D}{n}$. Consequently, the banker is disintermediated whenever she tries to renegotiate down the depositors.[23]

Case 3. $x \geq \frac{\tilde{D}}{n}$. Suppose first the banker does not attempt to repay less: she offers to repay $\frac{D}{n}$ per depositor. As before, assume that exactly $k \in \{0, ..., n\}$ depositors run on the bank. The banker can fully repay the $n-k$ depositors who do not run whenever $xS(k) \geq (n-k)\frac{D}{n}$. Since $\beta \in (0,1)$ and $x \geq \frac{\tilde{D}}{n}$ we get for all $k \in \{0, ..., n\}$:

$$x \geq \frac{D}{\beta} = \frac{k}{n}\frac{D}{\beta} + \frac{n-k}{n}\frac{D}{\beta} > \frac{k}{n}\frac{D}{\beta} + \frac{n-k}{n}D$$

$$\Rightarrow \quad x\left(1 - k\frac{D}{\beta x n}\right) > (n-k)\frac{D}{n}$$

The above implies $xS(k) > (n-k)\frac{D}{n}$ for all k between 0 and n. The banker can repay in full all remaining depositors regardless of how many reject her offer. All depositors accept whenever the banker offers to repay $\frac{D}{n}$ per depositor. Finally, suppose the banker offers to

[23] If not all depositors who accept are promised a pro-rate share, we apply the same argument but by considering a depositor who was promised less than $\frac{xS(k^*)}{n-k^*}$.

repay less than the initially promised $\frac{D}{n}$ per depositor. Since $\frac{D}{\beta x n} \leq 1$, any depositor who rejects and seizes assets is made whole. But then each depositor is strictly better off rejecting and seizing assets worth $\frac{D}{n}$ rather than accepting and being paid less than $\frac{D}{n}$. The banker gets disintermediated. She thus optimally chooses not to renegotiate the depositors down. In the unique equilibrium, the total repayment to the depositors is D, the payoff to the capital investor is $\beta(x - D)$, and the banker's payoff is $x - D - \beta(x - D) > 0$.