THE JOURNAL OF FINANCE • VOL. LX, NO. 2 • APRIL 2005

Liquidity Shortages and Banking Crises

DOUGLAS W. DIAMOND and RAGHURAM G. RAJAN*

ABSTRACT

We show in this article that bank failures can be contagious. Unlike earlier work where contagion stems from depositor panics or contractual links between banks, we argue that bank failures can shrink the common pool of liquidity, creating, or exacerbating aggregate liquidity shortages. This could lead to a contagion of failures and a total meltdown of the system. Given the costs of a meltdown, there is a possible role for government intervention. Unfortunately, liquidity and solvency problems interact and can cause each other, making it hard to determine the cause of a crisis. We propose a robust sequence of intervention.

BANKS ARE OFTEN FOUND AT THE CENTER of systemic financial crises. The causes of these crises have long been debated. Do they stem from depositor panics or do they arise because bank balance sheets are exposed to common factors? If it is the latter, then is there any sense in which bank failures can be contagious? Can anything be done to prevent these crises or stop them from propagating once they begin?

Our purpose in this paper is to reexamine whether bank failures can start and propagate systemic crises even if we assume that depositors do not panic, that is, even if depositors' actions are coordinated and they do not run simply because they fear others will run. Our focus is not because we think that panics are empirically unimportant (the jury is still out on that). Instead, we show that contagion can stem from an underlying real general equilibrium problem rather than just from lack of coordination among depositors. From a policy perspective, attempts to reassure investors through confidence-building measures will not work if panic and lack of coordination are not the reasons for a crisis.

We begin in this paper with a model of a bank that we developed in Diamond and Rajan (2001). Assets in this model are illiquid—they cannot be borrowed against or sold for the full value of resources they will generate in their best use—because the human capital necessary to generate that full value cannot

*Diamond and Rajan are from the University of Chicago and NBER, and University of Chicago, IMF, and NBER, respectively, we thank John Cochrane, Isabel Gödde, Gary Gorton, Nobuhiro Kiyotaki, and an anonymous referee for very helpful comments on an earlier draft; Steve Ross and participants at the NBER Economic Fluctuations and Growth meetings for helpful suggestions; and Efram Benmelech for great research assistance. We are grateful for financial support from the National Science Foundation and the Center for Research in Security Prices at the University of Chicago. Rajan also thanks the George J. Stigler Center for the Study of the Economy and the State for financial support.

be committed to the asset. So, for example, projects are illiquid because the entrepreneur who sets up the project is the best person to run the project. Even if he wants to pay out all the future revenues he generates to a potential lender (or if he wants to sell the project), he cannot get full value because he cannot commit to using his skills on behalf of the lender. Anticipating the rents the entrepreneur will extract in the future for his skills, the lender will only be willing to lend (or pay for) a fraction of the project value.

Someone who learns of a good alternative use of the project assets can lend more to the project because he can use the threat to put the assets to alternative use to collect more from the entrepreneur. Call such a lender the banker, for reasons that will become clear soon. The problem, however, is that while the project becomes less illiquid as the banker acquires collection skills, his loan to the project (i.e., the banker's asset) becomes more illiquid: Potential lenders to the bank do not have the banker's skills. So how then can the banker finance the project if it needs more funds than he has? Has illiquidity (stemming from specific skills) simply been transferred from the real project to the project loan?

The solution in Diamond and Rajan (2001) is that the banker can commit to using his collection skills on behalf of investors (and not extracting a rent) by issuing demand deposits. No depositor will ever accept a lower amount than initially promised. Because of the first-come, first-serve nature of deposits, the depositor is always better off demanding immediate payment if he thinks others will accept. Thus, any attempt by the banker to bargain down deposits will be met by a bank run, which Diamond and Rajan show completely dissipates any rents the banker could get. Fearing such a run, the banker will always pay what he has promised to depositors, if he has the means to pay.

By issuing demand deposits in sufficient quantity, the banker can effectively tie his collection skills to the loans he has made and borrow what is required from unskilled depositors. Thus, the banker does not pass on illiquidity, but reduces it through a combination of his collection skills and demand deposits. Moreover, so long as there is no aggregate shortage of resources, the bank's ability to issue new demand deposits allows it to meet an uncertain depositor demand for resources. Thus, banks play a central role in funding potentially long-term projects while allowing depositors to withdraw when in need.

Unfortunately, the demandable claims that perform such a useful role also expose the banking system to the risk that there might be a mismatch between the economy-wide production of resources and the demands of depositors. Even if the mismatch is merely due to delay in production, and not because of any impairment of the long-term production possibilities of the economy, such an aggregate shortage of resources (also termed an aggregate real liquidity shortage) can be spread through bank failures to the entire banking system.

Let us understand why. To do that, let us specify the nature of projects in greater detail. Projects require an investment of resources upfront, and pay off a larger amount of resources, either at an interim date (*early*) or after considerable delay (*late*).[1] Projects are thus risky, but only in that the timing of when they pay off is uncertain. Projects can also be discontinued and restructured so that they immediately return a fraction of the resources originally invested. We will see that the bank makes short-term loans to increase its bargaining power over the borrowers. We now ask what happens if a number of projects are delayed. To sharpen the intuition, let us assume that all initial depositors have utility for resources only at the interim date.

If the number of delayed projects is not too large in the economy, the banking system intermediates well. At the interim date, the bank can raise additional resources against loans to delayed projects and pass that along to depositors who demand them. The resources are available from entrepreneurs whose projects are early because they have spare resources after repaying their loans to their bankers. The bank's ability to collect from late projects and commit to repay these new depositors enables it to raise the resources, which it uses to repay old depositors. Thus, the economy-wide resources produced at the interim date are allocated efficiently via the bank.

Problems arise if too many projects are delayed in the economy so that there are too few resources produced at the interim date relative to depositor demand. Such an incipient aggregate shortage of fungible resources (i.e., liquidity) will draw forth three possible responses from bankers who are short of resources. First, they increase the real interest rate they are willing to pay for fresh deposits of resources. Since there is an aggregate shortage, this price alone is not enough to clear the market. Second, they attempt to sell late project loans. But since only the banker has the skills to collect on them, loan sales will not fetch much and this option could be dominated by the first one (of borrowing fresh resources to support the late loans). The third option is to call loans, forcing the borrowers to discontinue late projects, and restructure them to harvest resources immediately. Unlike the first two options, this option does increase the pool of resources available to satisfy the demand for liquidity, but it also reduces long-run production possibilities in the economy. As the real interest rate mounts, this option becomes more favored.

Unfortunately, the banks' asset values also drop with an increase in the real rate. Before banks collectively have enough incentive to produce the desired amount of aggregate liquidity by restructuring late projects, some banks may become insolvent. So a system-wide shortage of resources (a shortage of aggregate liquidity) at the interim date because of a delay in production can make banks insolvent.

More interesting is the converse. Insolvency can exacerbate liquidity shortages. The anticipated insolvency of a bank will precipitate an immediate bank run: Depositors will demand resources immediately rather than at the interim date. Faced with an immediate need for resources, the banker may prefer to

[1] This could be thought of as a model of delays in harvest, or crop failure, that so plagued banking systems in the nineteenth century. If the consumption good is commodity money or foreign exchange, the delay could be thought of as a delay of exports. It also has more modern parallels, which we will draw later.

restructure all projects (and not just the late ones) rather than to sell the illiquid project loans (we will make the conditions for this precise). But because early projects would have been produced at the interim date, which is the earliest anyone wants to consume, a bank run is socially detrimental. It forces the bank to produce resources immediately at great cost to the interim and long-run production possibilities of the economy. Moreover, because it subtracts from the pool of resources available at the interim date, a bank run can exacerbate aggregate liquidity shortages, making more banks insolvent, causing further runs, and squeezing the pool of resources even further till the entire system melts down. Problems in the banking system change real production decisions of borrowers, and this has economy-wide effects. This is an alternative view of contagious bank failures to that of Friedman and Schwartz (1963), who view it as working exclusively through the supply of money. Diamond and Rajan (2003) contains a more detailed comparison of our approach to that of Friedman and Schwartz.

Our approach suggests that the reason bank failures are contagious is also the same reason that bank assets are illiquid: Banks make hard-to-collect relationship loans and can influence the production decisions of their borrowers. Unfortunately, even though runs are socially inefficient, we argue that the right intervention by a central authority is by no means obvious, and that traditional interventions can be counterproductive.

Our work is related to a number of important recent papers. In particular, while Caballero and Krishnamurthy (2001, 2004) and Holmström and Tirole (1997, 1998) also focus on crises, the notion that banks create liquidity for depositors is less central to their arguments. Moreover, by liquidity they typically mean collateral value or wealth, while we introduce an additional notion of immediacy in this paper. The notion of loans as being illiquid due to the lender's specific lending skills, as well as his inability to commit, is in our earlier work (Diamond and Rajan (2001) as well as in Kiyotaki and Moore (2000)). Kiyotaki and Moore do not focus on the structure of bank contracts or on banking crises, while in Diamond and Rajan we examine bank deposit structures as commitment devices, assuming that there is a plentiful supply of aggregate liquidity.

Our paper is also related to Allen and Gale (2000), Bhattacharya and Gale (1987), Diamond (1997), Diamond and Dybvig (1983), Donaldson (1992), and Smith (1991), in its focus on the consequences of an aggregate shortage of liquidity. But there are three important differences between our paper and much of this literature. First, 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.

Second, we examine liquidity shortages stemming from the bank's asset side—from exogenous delays in the generation of project cash flows by borrowers or endogenous shortages stemming from the early termination of projects. Most of these earlier studies view liquidity shortages as stemming from the bank's liability side—from endogenous panics or exogenous fluctuations in the liquidity demanded by depositors. The difference is important. Individual bank runs in those models lead, at best, to fire sales of long-term assets, and at worst, to the premature liquidation of long-term assets if these assets are assumed to be nontransferable (as, e.g., in Allen and Gale (2000)). But runs then do not exacerbate shortages, since the run bank liquidates assets to pay off depositors. In fact, if the bank liquidates long-term assets to pay off depositors who do not need to consume, then a run can alleviate shortages, since these depositors can redeposit in sound banks (a flight to quality). Thus, if there is no aggregate liquidity shortage to begin with in these models, the mere fact that some banks become insolvent and are run cannot create a shortage. In our framework, by contrast, an aggregate liquidity shortage (and possible contagion) can be caused by bank runs precipitated by bank insolvency, even if there is no shortage in the absence of the runs.

The third difference stems from how contagion arises. In both Smith (1991) and Allen and Gale (2000), banks are linked ex ante, in the former through a banker's bank where investments are pooled, and in the latter through interbank loans. When the realized liquidity demand exceeds the supply, linked banks have to fail. In our paper, however, contagion can occur even if there are no explicit links between banks ex ante because of the negative real spillover effect of bank failure on the available liquidity. Banks are linked by a common market for liquidity.

The rest of the paper is as follows. We lay out the framework in Section I, and solve the model in Section II. Section III examines the effects of various kinds of interventions, after which we conclude.

I. Framework

A. Agents, Preferences, Endowments, and Technology

Consider an economy with three types of agents: investors, entrepreneurs, and bankers and four dates; 0, $\frac{1}{2}$, 1, and 2. Investors only get utility from near-term consumption, that is, their utility is a linear function of the consumption on or before date 1. All other agents also get equal utility from long-term consumption, and their utility is a linear function of the sum of consumptions at all dates on or before date 2. These assumptions are a very simple way of capturing the idea that some initial investors need to consume quickly. We discuss date $\frac{1}{2}$ below: it is a date when uncertainly is resolved and on which there may be renegotiation.

Investors are each initially endowed with a fraction of a unit of good. No other agent is endowed with goods. Goods can be stored at a gross real return of 1. They can also be invested in projects.

Each entrepreneur has a project that requires the investment of a unit of good at date 0. It pays C production goods at date 1 if the project produces *early* or pays C at date 2 if the project is delayed and produces *late*. There is a shortage of endowments of goods initially relative to projects that can be invested in, implying that entrepreneurs and banks must offer as high an expected return as possible to attract funding. There is a competitive market for bank deposits and other claims on banks.

B. Projects and the Nontransferability of Skills

The primary friction in the model is that those with specific skills earn a rent from future surplus produced because they cannot commit to using their human capital on behalf of others. This implies that they are not able to borrow the full value of the surplus they can produce with an asset or sell the asset for the amount they can produce with it. Both projects and loans to projects are illiquid because of the inalienability of human capital.

Specifically, since an entrepreneur has no endowment, he needs to borrow to invest. The entrepreneur can, however, credibly threaten not to supply his human capital (destroying the project if there isn't outside intervention) at any date, including date $\frac{1}{2}$. This credible threat would allow him to renegotiate down any promised payment, typically making it impossible for the entrepreneur to borrow. Fortunately, he has access to a banker who has (or who can acquire during the course of lending) knowledge about an alternative, but less effective way to run the project. If the banker lends, while keeping the right to foreclose on the project any time the entrepreneur threatens to withdraw, he can counter the entrepreneur's threat with the credible threat of seizing assets. Thus, the banker's specific knowledge, together with the demandability of the loan he makes, allows the banker to collect γC from an entrepreneur whose project just matures.[2] No one else has the knowledge to collect from the entrepreneur.

Regardless of whether a project is early or late, the banker can also *restructure* the project at any time before date 1 to yield c_1 in goods immediately and c_2 in date-2 goods—intuitively, restructuring implies stopping half-finished projects, and salvaging all possible production goods from them (equivalently, the banker can call in the loan and force the borrower to restructure). Restructured projects can be collected by anyone. We assume

$$c_1 + c_2 < 1 < \gamma C < C, \tag{1}$$

Since no one other than the bank has the specific skills to collect from the entrepreneur, the loan to the entrepreneur is illiquid, in that the banker will get less than γC if he has to sell the loan before the project matures. Any buyer realizes that the banker extracts a future rent for collecting the loan, and the buyer reduces the price he pays for the loan accordingly. In fact, bank loans are so dependent on the banker's specific skills for collection (i.e., they are so illiquid) that the banker prefers restructuring projects to selling them.[3]

Since there is a shortage of endowment relative to projects, only a select few entrepreneurs get a loan from their respective banks to buy a unit of the good from investors. Entrepreneurs have to promise to repay the maximum possible to obtain the loan. This means that the amount the borrower must promise to pay the banker on demand is γC.

[2] While we do not derive this here, see Hart and Moore (1994) or Diamond and Rajan (2001) for an extensive form game with this outcome.

[3] This requires $c_1 + c_2 \geq \gamma C - \pi$, where π is the discounted present value of the banker's rent.

C. Financing Banks

We analyze the general equilibrium effects in an economy in which banks finance the illiquid loans with both demand deposits and bank capital (which in our model is any claim that is long-term, and cannot demand payment on date $\frac{1}{2}$, including equity, and longer term debt). Our positive results do not depend on the reason that banks choose this capital structure. However, in previous work, we argue that bankers can commit to collect loans on behalf of investors if they issue demand deposits to them. It is useful to briefly recount the reasons here.

Bankers themselves have no endowment, so they have to persuade investors to entrust them with their goods. But unlike the banker, ordinary investors cannot collect from the entrepreneur. The problem then is that having obtained investors' money promising a certain repayment, the banker can threaten to hold back his collection skills unless investors reduce the required repayment. Since courts cannot compel the banker to contribute his human capital even if courts can enforce financial contracts, investors are vulnerable to attempts at strategic renegotiation by the banker. The prospect of having the promised repayment renegotiated down would seriously impair the amount that investors are willing to entrust the banker with. Therefore, the banker has to find a way to commit to using his skills on behalf of investors, or else he will not be able to raise enough to finance the loan he has made.

The banker can finance lending while committing his human capital to the service of investors by issuing uninsured demand deposits. Because of the first-come, first-served aspect of uninsured demand deposits, they cannot be negotiated down. This is because depositors are liable to run to demand repayment if they ever apprehend that they will be paid less than their due. The run is in the individual interest of each depositor (it is a Nash equilibrium). The run forecloses on the bank's assets even though it may give depositors less collectively than if they negotiated to accept a payment lower than was promised originally. Courts enforce the depositors' demands, so the banker has to pay, and this destroys the banker's rents (see Diamond and Rajan (2001) for details). Thus, if a banker has promised to pay depositors d_t, they want to consume at date t, and if the banker has enough resources at that date, he makes the payment rather than risk a run by renegotiating.[4]

In Diamond and Rajan (2001) there is no uncertainty, and fixed-value demand deposits are optimal contracts because runs are always off the equilibrium path. For these contacts to be optimal here, we need to assume that the amount promised to each demand depositor cannot be made contingent on the bank specific or aggregate state of nature. Alternatively, our analysis can be seen as an examination of the positive effects of using uncontingent demand deposits.

[4] For other models where runs or short-term debt serve as a source of discipline, see Calomiris and Kahn (1991) and Jeanne (2000). The difference in Diamond and Rajan (2001) is that the run is particularly useful in disciplining an intermediary, even if it is not effective in disciplining a corporation that borrows directly.

Since the banker can credibly threaten to not collect the loan at any point after the deposit is made, deposits must be demandable at any time if they are to provide commitment, although consumption is needed only on dates 1 or 2. In fact, demandability is critical at date $\frac{1}{2}$ (a stand-in for any time between dates 0 and 1) even though it hurts depositors collectively, if indeed they demand payment at this earlier time (as we will see). Therefore, the nature of the bank's lending activities creates a necessary fragility in its capital structure—deposits are demandable, not so much to satisfy the uncertain consumption needs of depositors (though, as we will see, they end up doing that) but so that the banker can commit, via the collective action problem created by deposits, not to extract rents from depositors.[5]

But while the collective action problem inherent in demand deposits enables the banker to commit to repay if he can (i.e., to avoid strategic defaults), it exposes the bank to destructive runs if he truly cannot pay (it makes nonstrategic default more costly): When depositors demand repayment before projects have matured, the bank is forced to restructure projects to get c_1 immediately instead of allowing them to mature and generate γC.

Such rigidity can be tempered by raising part of the funds through claims that can be renegotiated. In particular, we focus on the issue of capital where the renegotiation process leads the banker and capital holders to split the residual surplus after deposits have been paid (see the extensive form game in Diamond and Rajan (2000)). Assuming equal division, capital will be paid $(v-d)/2$, where v is the present value of the bank's assets in their best use from capital's perspective and d is the level of deposit repayments. So the virtue of capital is that its value adjusts to the underlying value of the bank's assets; its cost is that the bank cannot raise as much upfront against its future receipts, since the banker absorbs an equal surplus that increases in the level of capital. We assume that capital has to be fraction k of the present value of bank assets (this could be because of unmodeled uncertainty the bank faces on its loans or because of an explicit capital requirement imposed by regulators) so that the bank can raise only

$$\frac{v}{1+k} \tag{2}$$

against its assets by issuing a mix of deposits and capital.[6] We add bank capital to the model only to understand its effects: The key qualitative results are

[5] We assume a court system in the background, which enforces contracts and transfers assets to lenders when loans are defaulted upon. The court, however, cannot force agents to supply their human capital. So if the banker borrows from a single lender, the banker can threaten to not provide his collection skills. The lender can use the court to seize the bank's loans, but these are of little value to the lender. So the lender will be renegotiated down. By contrast, the collective action problem inherent in demand deposits forces a change in the ownership of the bank's assets whenever the bank attempts to renegotiate. As Diamond and Rajan (2001) show, this becomes the source of discipline.

[6] From the definition, we have $k = [\frac{1}{2}(v-d)/\frac{1}{2}(v+d)]$ where the numerator on the right-hand side is the value of capital and the denominator is the value of capital plus maturing deposits. Therefore, the total amount that can be pledged to investors out of the amount the bank collects is the denominator, which on substituting for d works out to $v/(1+k)$.

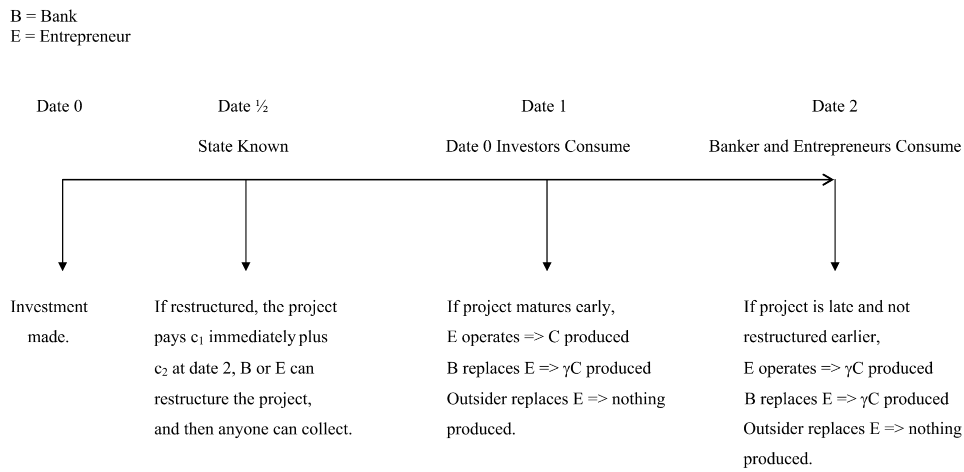

Figure 1. Project outcomes.

unchanged if we leave it out and assume that all claims on the bank are demand deposits.

D. Sequence of Events in the Economy

Each bank faces an identical pool of entrepreneurs at date 0. At date $\frac{1}{2}$, everyone learns which projects are early and which are late. The aggregate state is then fully characterized by the realization of the fraction of projects that are early for each bank i, α^i. Without loss of generality, let us assume that α^i is increasing in the index i. The realized distribution of banks in the aggregate state s is $F^s(\alpha^i)$. The date-0 probability of the aggregate state s is q^s.

In what follows, we suppress the dependence on the state for notational convenience. All quantities are henceforth normalized by the total initial endowment of goods. The timeline for a project and the loan to it is in Figure 1. The sequence of events is summarized in Figure 2. Ex ante identical banks compete for investors' endowments by issuing a mix of deposits d_0 (per unit of consumption good invested), and given the distribution of anticipated outcomes, the associated quantity of capital to them at date 0. Since consumption goods are in short supply relative to projects, investors are promised the maximum that can be pledged to them and invest so long as their opportunity rate of return (of storage) is met.

The bank can then lend to entrepreneurs or invest in storage. We show later that the bank does not store if it lends, but we assume it for now. If it lends, the bank will charge γC repayable on demand.

At date $\frac{1}{2}$, everyone learns what fraction α^i of a bank i's projects are early. If depositors anticipate that the bank will not be able to pay them at date 1, given the bank's state and aggregate liquidity conditions, they will run immediately; otherwise they will wait till date 1 and demand payment in an orderly fashion. Runs are, therefore, based on fundamentals: Because we wish to understand

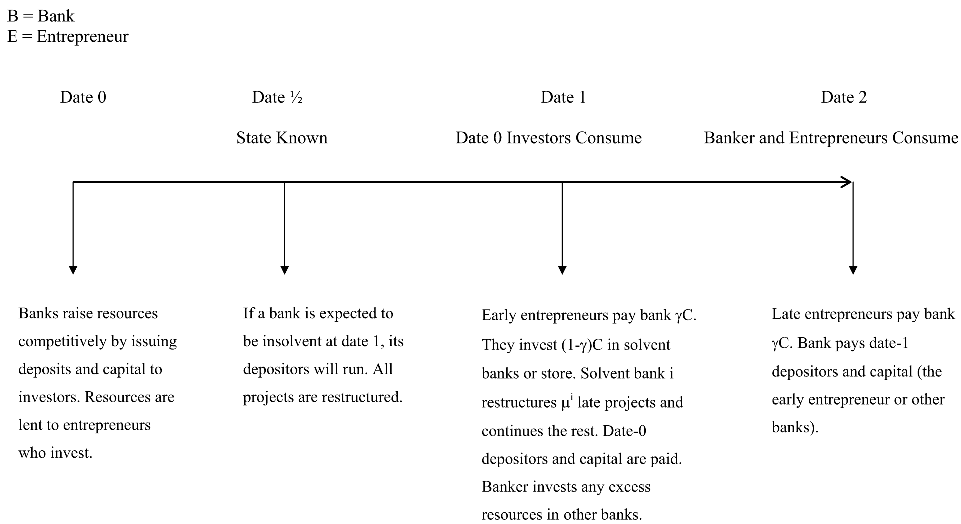

Figure 2. Sequence of events.

how aggregate liquidity shortages influence bank decisions and possibly lead to contagious bank failures, we do not allow runs caused by panics, where depositors run at date $\frac{1}{2}$ only because they think other depositors will run, regardless of date-1 fundamentals (see, e.g., Diamond and Dybvig (1983)). We also assume for now that if there is more than one market-clearing interest rate, depositors can coordinate their beliefs about this rate.[7]

If their bank survives, the entrepreneurs with early projects repay the bank γC at date 1 (leaving them with $(1 - \gamma)C$ to invest as they will), while entrepreneurs with late projects default. Depending on the prevailing rate and on the bank's need for funds, the bank then decides how to deal with each late project—whether to restructure it if goods are needed at date 1, or perhaps get greater long-run value by rescheduling the loan payment till date 2 and keeping the project as a going concern. The bank uses repayments from entrepreneurs whose projects are early, from restructured late projects, and from new funds reinvested in bank deposits and capital by early entrepreneurs to repay investors at date 1. Finally, at date 2, late projects mature and their proceeds are divided up according to commitments made at date 1. Bankers and entrepreneurs consume.

II. Solving the Model

We determine individual optimizing decisions, then aggregate them, applying market-clearing conditions to obtain the equilibrium. The optimizing decisions

[7] In other words, while we assume that each depositor expects the others in the same bank to choose a withdrawal that is an individual best response to others' actions (so we assume noncooperative actions where individual incentives may not lead each depositor to maximize the welfare of the whole), they all agree to choose the set of Nash actions that makes them best off.

of all agents in this model except the bankers are relatively simple. Investors deposit in banks provided that their expected gross return from date 0 to date 1 weakly exceeds storage. They wait till date 1 to withdraw, unless they know that the bank cannot repay everyone, in which case they run as soon as uncertainty is revealed at date $\frac{1}{2}$. If their projects are not restructured earlier, early entrepreneurs produce at date 1 and repay what they owe. They invest the excess provided that the real rate paid on bank claims between date 1 and date 2, $r \geq 1$. Late entrepreneurs produce at date 2 if their projects are not restructured earlier, and also repay what they owe.

A. Banker's Maximization

The banker's decision is only a little more complicated. Taking the real interest rate between date 1 and date 2, r, as given, banker i has to decide the fraction of late projects to restructure, μ^i, so as to maximize his consumption while constrained by the necessity to pay off all bank claimants. This problem can be written as

$$\max_{\mu^i} \alpha^i \gamma C + \mu^i (1-\alpha^i)\left(c_1 + \frac{c_2}{r}\right) + (1-\mu^i)(1-\alpha^i)\frac{\gamma C}{r}, \tag{3}$$

$$\text{s.t. } v(\alpha^i, \mu^i, r) \geq \frac{\max\limits_{\mu'} \;\; v\left(\alpha^i, \mu', r\right) + d_0}{2}. \tag{4}$$

To see why this is his maximization problem, start with the constraint (4). The value the banker can raise at date 1 in consumption goods is

$$v(\alpha^i, \mu^i, r) = \alpha^i \gamma C + \mu^i (1-\alpha^i)\left(c_1 + \frac{c_2}{r}\right) + (1-\mu^i)(1-\alpha^i)\frac{\gamma C}{(1+k)r}. \tag{5}$$

The first term is the amount repaid by the α^i early entrepreneurs whose projects mature at date 1. The second term is the amount obtained by restructuring late projects. The third term is the amount the bank can raise (in new deposits and capital—see (2)) against late projects that are allowed to continue without interruption till date 2.

In bargaining about how much they have to be paid, capital holders have the right to pick a level of restructuring that maximizes what they get.[8] This can be different from the level actually chosen by the banker. Hence the repayment to capital and deposits is $[\max_{\mu'} v(\alpha^i, \mu', r) + d_0]/2$. Therefore (4), the constraint on the banker, is simply that the resources he can raise should exceed the required repayment.

The banker's objective is to maximize the present value of his total consumption. Since the bank's repayment to initial investors is invariant to the actual amount of restructuring chosen (repayment depends instead on the amount

[8] Intuitively, if capital has the right to make a take-it-or-leave-it offer to the banker, it demands all the value the banker generates net of deposits when the assets are put to their best use *from the perspective of capital*.

of restructuring that maximizes outsiders' claims), the banker is the residual claimant of repayments by entrepreneurs on bank loans, and his objective can be changed to maximize the present value of the total repayments on his loans, as in (3). Note that this exceeds (5), the amount he can raise at date 2 because the banker cannot raise money against his own prospective future rents.

The banker's decision boils down to whether he should continue or restructure a late project. To see the solution to this problem, note that the banker can raise $\gamma C/(1+k)r$ in deposits and capital at date 1 against the prospective payment from a late entrepreneur. He obtains $c_1 + (c_2/r)$ by restructuring the late project. So continuation (or $\mu^i = 0$) maximizes the value the bank raises, $v(\alpha^i, \mu^i, r)$, and thus loosens the constraint (4) most when $r < R = [\gamma C/(1+k) - c_2]/c_1$, while restructuring ($\mu^i = 1$) loosens it most when $r \geq R$. By contrast, the banker himself cares about the value he gets (including his rents) at date 2. If he continues a late project, he collects γC at date 2. If he restructures it, he gets $rc_1 + c_2$ at date 2. Therefore, the banker's objective function is maximized by continuing a project if $r < \bar{R} = (\gamma C - c_2)/c_1$ and restructuring it otherwise. It follows then that the solution to the constrained maximization is $\mu^i = 0$ if $r < R$ and $\mu^i = 1$ if $r \geq \bar{R}$. But for moderate rates such that $R \leq r < \bar{R}$ because the banker earns an unpledgeable rent on late projects, he wants to continue them, even though he can raise more at date 1 by restructuring them. It follows that the banker sets μ^i at the minimum level that satisfies the constraint. So

LEMMA 1: *(i) When $1 \leq r < R$, no solvent bank restructures late projects. Further, let $i^*(r)$ be such that $v(\alpha^{i^*}, 0, r) = d_0$. Then banks with $i \geq i^*(r)$ are solvent, while those with $i < i^*(r)$ are insolvent and are run. (ii) When $r \geq \bar{R}$, solvent bank i restructures all its late projects. A bank is solvent if $i \geq i^*(r)$ where $i^*(r)$ is such that $v(\alpha^{i^*}, 1, r) = d_0$. (iii) When $R < r < \bar{R}$, solvent bank i restructures a fraction of late projects, which is the minimum nonnegative μ^i such that $v(\alpha^i, \mu^i, r) \geq [v(\alpha^i, 1, r) + d_0]/2$. A bank is solvent if $i \geq i^*(r)$ where $i^*(r)$ is such that $v(\alpha^{i^*}, 1, r) = d_0$.*

Proof: Omitted.

Lemma 1 suggests that, ceteris paribus, the higher the prevailing interest rate, the more attractive it becomes for a bank to restructure its late projects. Also, as the interest rate goes up, the bank's asset value falls, which could lead the bank to become insolvent and be run.

As soon as uncertainty is revealed at date $\frac{1}{2}$, a bank's depositors estimate whether the bank will be solvent at date 1 at the predicted equilibrium interest rate—whether there is some nonnegative μ^i such that (4) is satisfied—or run immediately if they anticipate it will not. Even though depositors have no need to consume before date 1, the anticipated demise of the bank forces them to rush for payment before others. Since bank assets are illiquid and cannot be transferred without restructuring, and since the run occurs before date 1, the bank restructures all projects, even early ones, in a futile attempt to pay off depositors as they come to withdraw.[9] The bank produces liquidity in a very inefficient way, restructuring early projects to produce immediate liquidity, even though no one can consume before date 1. Depositors, however, leave it with no alternative. The value paid out to depositors is the entire market value of the bank's assets, $c_1 + (c_2/r)$, and all projects are restructured.

Interestingly, even if early projects in run banks are being restructured to produce liquidity inefficiently, not all banks have the incentive to restructure late projects and add to the available supply of liquidity. The reason is that because the run bank is insolvent (and cannot afford to pay the social value for retaining deposits), the interest rate does not reflect the value of liquidity to the banking system. In fact, if $r < R$, Lemma 1 (i) indicates that the solvent bank continues all late projects, regardless of whether other banks are being run.

B. Market-Clearing Conditions

Given the initial level of deposits, the equilibrium is one where all agents maximize given the state, and the goods market clears at date $\frac{1}{2}$, date 1, and date 2. Because date-$\frac{1}{2}$ consumption is a perfect substitute for date-1 consumption (and goods can also be stored), the supply of goods on those two dates can be added together to meet date-1 claims. The demand for consumption before date 2 comes entirely from initial investors who can express their demand through bank claims. For a solvent bank, the value of bank claims, and hence demand, is $[\max_{\mu'} v(\alpha^i, \mu', r) + d_0]/2$. For an insolvent bank, it is $c_1 + (c_2/r)$. The supply comes from early projects and restructured projects. A solvent bank i supplies $l(\alpha^i, \mu^i) = \alpha^i C + \mu^i(1 - \alpha^i)c_1$, while an insolvent bank supplies c_1. The real interest rate clears the market for consumption before date 2 (i.e., liquidity). If this market clears, the date-2 goods market also clears automatically by Walras law.

In what follows, we will see that there can be multiple interest rates consistent with the liquidity market clearing. Let us term the equilibrium where each investor's belief about date-1 interest rate is the lowest consistent with market clearing the "most optimistic equilibrium." We call it this because initial investors all wish to consume at date 1, and each is better off if their banks do not fail and if real interest rates prevailing when their banks are attracting new

[9] When a bank is run and it runs out of reserves, it can either allow depositors to seize project loans equivalent in market value to what they are owed, or restructure projects and pay the resulting funds out to depositors (by assumption of illiquidity, restructuring dominates the bank selling the loan). In either case, the bank's rents are zero, but the bank has a weak preference for restructuring rather than allowing depositors to seize loans. The reason is that the market value of seized loans is the value from restructuring them (the best alternative use of the project assets) less the strictly positive rent the bank is anticipated to charge for restructuring. Thus the market value of loans is less than the value from restructuring the underlying projects. Therefore, the banker weakly prefers restructuring loans when faced with a run, and if there is any chance that the queue of depositors will not exhaust the bank's assets, this preference is strict. Thus, when run, the banker restructures all projects.

deposits to repay them are low (this also increases the market value of their long-term [capital] claims). As a result, they are always worse off coordinating on beliefs that artificially increase future interest rates. It is particularly useful to show that contagion is possible when we rule out panics and beliefs coordinated on the "wrong" interest rates—that is, even if we focus only on the most optimistic equilibrium. We do, however, comment later on other possible equilibria when there exist multiple Bayesian Nash equilibria.

Now consider the market for liquidity. If there is an excess of supply over demand at an interest rate of 1 (the storage rate), that is the real rate and the excess goods are stored. By contrast, if there is excess demand at a rate of 1, the real interest rate has to rise. It is easy to show that the interest rate in the most optimistic equilibrium is given by

PROPOSITION 1:

(i) $r = 1$ *if*

$$\int_{\underline{i}}^{i^*(1)} c_1 f(i)\,di + \int_{i^*(1)}^{\bar{i}} l(\alpha^i, 0) f(i)\,di \geq \int_{\underline{i}}^{i^*(1)} [c_1 + c_2] f(i)\,di + \int_{i^*(1)}^{\bar{i}} \frac{1}{2}\big[v(\alpha^i, 0, 1) + d_0\big] f(i)\,di \tag{6}$$

(ii) *r greater than* 1 *and less than R if* (6) *does not hold but*

$$\int_{\underline{i}}^{i^*(R)} c_1 f(i)\,di + \int_{i^*(R)}^{\bar{i}} l(\alpha^i, 0) f(i)\,di \geq \int_{\underline{i}}^{i^*(R)} \left[c_1 + \frac{c_2}{R}\right] f(i)\,di + \int_{i^*(R)}^{\bar{i}} \frac{1}{2}\big[v(\alpha^i, 0, R) + d_0\big] f(i)\,di \tag{7}$$

(iii) *r greater than or equal to R and less than* $\bar{R}$ *if* (6) *and* (7) *do not hold but*

$$\int_{\underline{i}}^{i^*(\bar{R})} c_1 f(i)\,di + \int_{i^*(\bar{R})}^{\bar{i}} l(\alpha^i, \mu^i) f(i)\,di > \int_{\underline{i}}^{i^*(\bar{R})} \left[c_1 + \frac{c_2}{\bar{R}}\right] f(i)\,di + \int_{i^*(\bar{R})}^{\bar{i}} \frac{1}{2}\big[v(\alpha^i, 1, \bar{R}) + d_0\big] f(i)\,di \tag{8}$$

(iv) $r \geq \bar{R}$ *otherwise where r solves*

$$\int_{\underline{i}}^{i^*(r)} c_1 f(i)\,di + \int_{i^*(r)}^{\bar{i}} l(\alpha^i, 1) f(i)\,di = \int_{\underline{i}}^{i^*(r)} \left[c_1 + \frac{c_2}{r}\right] f(i)\,di + \int_{i^*(r)}^{\bar{i}} \frac{1}{2}\big[v(\alpha^i, 1, r) + d_0\big] f(i)\,di \tag{9}$$

Proof: Omitted.

One might think that a higher interest rate (effectively, the price of liquidity) would tend to reduce the excess demand for liquidity. It turns out that this is not always true because of the effect of bank failures. To see why, consider first the reasons why it would reduce the excess demand. For example, when $1 \leq r < R$ so that solvent banks continue all late projects, the excess demand is:

$$ED_{r<R} = \int_{\underline{i}}^{i^*(r)} \left\{ \left[c_1 + \frac{c_2}{r} \right] - c_1 \right\} f(i)\, di + \int_{i^*(r)}^{\bar{i}} \left\{ \frac{1}{2} \left[\alpha^i \gamma C + \frac{(1-\alpha^i)\gamma C}{(1+k)r} + d_0 \right] - \alpha^i C \right\} f(i)\, di. \qquad (10)$$

The term in the first set of curly brackets is the excess demand from banks that are insolvent, while the term in the second set is the excess demand from solvent banks. Note that even though insolvent banks restructure all their loans, they also sell the date-2 portion of restructured loans for goods. On net, therefore, they subtract liquidity of c_2/r from the overall pool. Differentiating the excess demand with respect to r, we get

$$\frac{dED_{r<R}}{dr} = -\int_{\underline{i}}^{i^*(r)} \frac{c_2}{r^2} f(i)\, di - \int_{i^*(r)}^{\bar{i}} \frac{1}{2} \frac{(1-\alpha^i)\gamma C}{(1+k)r^2} f(i)\, di + \frac{di^*}{dr} f(i^*) \left[\frac{c_2}{r} - \left[\frac{(1-\alpha^{i^*})\gamma C}{(1+k)r} - \alpha^{i^*}(1-\gamma)C \right] \right]. \qquad (11)$$

A higher interest rate reduces demand because depositors in failed banks just get the bank's asset value, and this (in particular, the date-2 portion of restructured loans) falls with the interest rate. Hence the first term in (11). While depositors in solvent banks get an amount that is invariant in the interest rate, the holders of capital in solvent banks are paid less because bank asset values (in particular, the value of continued loans) fall with the interest rate. So demand falls for this reason also, hence the second term in (11). For $r < R$, solvent banks do not restructure loans. For $r \geq R$, they do, so an additional term would be the increase in supply (and the fall in excess demand) from the additional restructuring forced by a rise in the interest rate. Thus, if one were to ignore the effect of bank failures on excess demand (the last term in (11)), excess demand would always fall with the interest rate.

C. *Example 1*

Consider a realized state of nature with two possible types of banks at date 1: L and H, where the banks are distinguished by the fraction of early projects in their portfolio, either $\alpha^{\mathrm{L}} = 0.55$ or $\alpha^{\mathrm{H}} = 0.7$. In the realized state, let fraction 0.2 of banks be of type L and 0.8 of banks be of type H. Let $c_1 = 0.4, c_2 = 0.5, C = 1.6, \gamma = 0.8, k = 0.12$. Plugging in values, we obtain $\bar{R} = 1.95$ and $R = 1.60$.

Let the level of outstanding deposits per unit loaned at date 0 be $d_0 = 1$. When discussing aggregates, we normalize the total amount loaned out in the system to 1.

In this example, there is sufficient date 1 liquidity in the most optimistic equilibrium to repay all deposits without any restructuring ($0.2\alpha^{\mathrm{L}}C + 0.8\alpha^{\mathrm{H}}C = 1.072 > 1 = d_0$). However, at a real interest rate equal to that of storage, $r = 1$, there is an excess demand for liquidity because capital also has to be paid dividends which, aggregated across banks, exceed 0.072. Higher interest rates reduce the value of capital's dividends, until at $r = 1.31$, the aggregate value of capital dividends plus deposits is 1.072; $r = 1.31$ clears the market without any project restructuring (because $r < R = 1.60$). For higher levels of α^{L}, there is more date-1 aggregate liquidity. If we change α^{L}, keeping $\alpha^{\mathrm{H}} = 0.7$ and the fraction of type L banks constant, the equilibrium interest rate, r, declines monotonically to 1 as α^{L} increases from 0.55 to 0.7. Eventually, with $\alpha^{\mathrm{L}} \geq 0.7$, there is more than enough date-1 liquidity to pay the dividends to capital even with $r = 1$, the real return of storage.

Now suppose the fraction of type L banks increases to 0.6. Let us vary the realized fraction of early projects of type L banks, leaving everything else unchanged. Now we get restructuring. For example, with $\alpha^{\mathrm{L}} = 0.51$ there is a liquidity shortage if no bank restructures late projects because aggregate date 1 liquidity is insufficient to repay all deposits: $[0.6\alpha^{\mathrm{L}} + 0.4\alpha^{\mathrm{H}}](1.6) = 0.9376 < 1 = d_0$. The market clearing interest rate is $r = 1.61$ in this case, which gives L banks the incentive to restructure projects (since the interest rate exceeds $R = 1.60$) while leaving them solvent (they fail only if $r > 1.94$). If we change α^{L}, the equilibrium interest rate is:

$\alpha^{\mathrm{L}} = 0.51$	$R = 1.61$	L banks restructure all late projects (L just solvent)
$\alpha^{\mathrm{L}} = 0.55$	$R = 1.84$	L banks restructure some late projects (L just solvent)
$\alpha^{\mathrm{L}} = 0.56$	$R = 1.87$	L banks restructure some late projects (L just solvent)
$\alpha^{\mathrm{L}} = 0.59$	$R = 1.91$	L banks restructure no late projects (L just solvent)
$\alpha^{\mathrm{L}} = 0.6$	$R = 1.70$	No banks restructure (L capital has value)
$\alpha^{\mathrm{L}} = 0.65$	$R = 1.31$	No banks restructure (L capital has value)
$\alpha^{\mathrm{L}} = 0.7$	$R = 1$	No banks restructure (L capital has value)

Interestingly, the real interest rate initially goes up as the fraction of early projects in the L type banks increases. The increase in α^{L} from 0.51 to 0.59 increases the L type bank's intrinsic ability to pay higher interest rates. So long as the alternative is to restructure projects, each type L bank promises the highest interest rate consistent with its solvency (because banks prefer to offer higher interest rates rather than to restructure projects at real interest rates between R and $\bar{R}$). But when $\alpha^{\mathrm{L}} \geq 0.6$, there is sufficient liquidity for all banks to survive without any restructuring, and increases in α^{L} then reduce real interest rates as before. The interesting point is that there is not a monotonic relationship between the amount of aggregate liquidity absent restructuring, which increases in α^{L}, and the real rate.

D. Bank Failures and Contagion

Thus far we have ignored the effect of a rising real rate on bank failures. Now consider the third term in (11). An increase in the interest rate causes more banks to become insolvent and run. Whether this increases or decreases excess demand depends on the sign of the term in square brackets. It is the difference between the excess demand of an insolvent bank ($=c_2/r$) and the excess demand of a bank that is just solvent $[(1-\alpha^{i^*})\gamma C/(1+k)r] - \alpha^i(1-\gamma)C$ (the first term in this last expression is the liquidity that the solvent bank absorbs from the system by continuing late loans, and the second term is the liquidity it indirectly generates for the system by allowing early entrepreneurs with spare goods to produce). Bank failure reduces excess demand if the bank has few early projects (α low) so that it does not produce much liquidity while solvent but instead absorbs a lot to keep late projects afloat. Such a bank may even release liquidity on net in the process of being run because it will be forced to restructure its many late projects (in addition to the few early ones). So at interest rates where the marginal insolvent bank has low α, the overall excess demand for liquidity is likely to fall in the interest rate.

But a bank could well add to excess demand when it fails. For example, when $r > R$, the marginal solvent bank already liquidates all late projects, so failure does not create any additional liquidity. Rather, because bank failures cause even early projects to be restructured before they produce liquidity, failure increases the excess demand for liquidity.[10] It is now possible for the aggregate excess demand to increase with interest rates (all i.e., required for this is that the increase in excess demand from the mass of additional failing banks outweighs the reduction in excess demand as a result of higher rates). In this case, bank runs are destabilizing, and an initial shortage of liquidity could result in a meltdown of the entire banking system.

E. Example 2

Now let us return to the earlier example, but let us set $\alpha^{\mathrm{L}} = 0.4$ and $\alpha^{\mathrm{H}} = 0.6$, while a fraction 0.6 of the banks are of type L. In the absence of any restructuring, the aggregate supply of liquidity is only $0.6[\alpha^{\mathrm{L}}C] + 0.4[\alpha^{\mathrm{H}}C] = 0.768$, which is insufficient to pay deposits and capital. Faced with an aggregate excess demand for liquidity, banks bid up rates. The L type banks (with $\alpha^{\mathrm{L}} = 0.4$) are insolvent for $r > 1.40$. Since no bank voluntarily restructures to increase supply for any interest rate below $R = 1.60$, and since some additional supply (or a drastic reduction in demand) is needed, this means the equilibrium interest rate has to be above 1.60, and L type banks are insolvent in equilibrium.

[10] All the bank's late projects are restructured. So the bank's projects spew out $\alpha^i C + (1-\alpha^i)c_1$ of cash at date 1 if the bank stays solvent (note that $\alpha^i\,(1-\gamma)C$ of this stays with the early entrepreneur). Since the bank is just solvent, all the date-1 amount it can raise, $\alpha^i\gamma C + (1-\alpha^i)[c_1 + c_2/r]$, just meets what it owes date-1 claimants. So the net amount of liquidity the bank releases to the system when solvent is $\alpha^i\,(1-\gamma)C - (1-\alpha^i)c_2/r$. By contrast, the net amount of liquidity released when a bank restructures all projects is $c_1 - (c_1 + (c_2/r)) = -c_2/r$, which is clearly lower. So a bank failure absorbs more liquidity from the system relative to a solvent bank.

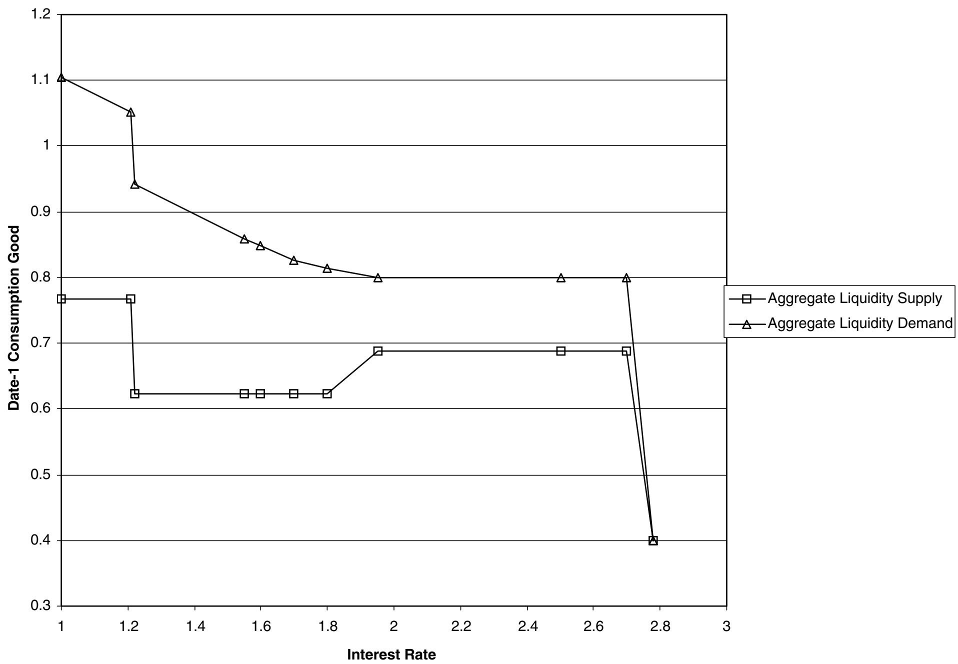

Figure 3. Example 2—liquidity supply and demand as functions of real rate, *r*.

Anticipating this, depositors of L type banks run at date $\frac{1}{2}$. Figure 3 shows the supply and demand for liquidity.

How does the failure of the L type banks change the aggregate excess demand for liquidity? On the one hand, the demand for liquidity falls because depositors in the failed banks have to settle for $c_1 + (c_2/r)$ instead of the 1 they are owed, and capital gets nothing. Thus, bank failure at $r = 1.6$ reduces demand by $d_0 - c_1 - (c_2/r) = 0.29$.[11] But a run also reduces the potential supply of liquidity. An L type bank could potentially create $\alpha^{\mathrm{L}}C + (1 - \alpha^{\mathrm{L}})c_1$ of liquidity. But when it is run at date $\frac{1}{2}$, it creates only c_1. Thus, there is a reduction in liquidity supply of $\alpha^{\mathrm{L}}C - \alpha^{\mathrm{L}}c_1 = 0.48$ per run bank. Thus, the failure of L type banks with aggregate weight 0.6 *increases* the aggregate excess demand by 0.6 (0.48–0.29) = 0.11.

Put another way, the depositors of the failed bank absorb all of the liquidity produced by their bank plus c_2/r. The net supply of liquidity (at $r = 1.60$) available to surviving H banks in aggregate is then $0.4[\alpha^{\mathrm{H}}C] - 0.6(c_2/r) = 0.197$, not enough to pay their deposits (which aggregate to 40% of the total, or 0.4). Thus, the interest rate has to rise further to get H banks to start restructuring projects. At $r = \bar{R} = 1.95$, the H banks are willing to restructure all their late projects. Even so, the aggregate supply of liquidity available to them is $0.4[\alpha^{\mathrm{H}}C + (1 - \alpha^{\mathrm{H}})c_1] - 0.6\frac{c_2}{r} = 0.448 - 0.6\frac{c_2}{r} = 0.294$. As a result, even after

[11] At the point of failure, capital gets nothing.

the type H banks liquidate all their late projects, there is still insufficient liquidity for them to repay their deposits of 0.4. Eventually, as the interest rate climbs still higher to 2.77, even type H banks also fail before demand falls enough to meet supply. Thus, all banks fail, and in aggregate, depositors consume $c_1 = 0.4$, while bankers and entrepreneurs consume $c_2 = 0.5$ (if depositors had placed any value on date-2 consumption, this would be consumed by them).

The total aggregate liquidity that could be generated if all late projects were restructured (but no banks are run) is $0.6[\alpha^{\mathrm{L}}C + (1 - \alpha^{\mathrm{L}})c_1] + 0.4[\alpha^{\mathrm{H}}C + (1 - \alpha^{\mathrm{H}})c_1] = 0.976$. Even when outstanding deposits are just 0.976 rather than 1—so that there is enough potential aggregate liquidity for both banks to survive—the only equilibrium is still that both banks fail. Here is why: The private incentive of each bank is to free-ride on publicly available liquidity until the price of liquidity becomes high enough $(=\bar{R})$ to make it worthwhile to supply some by restructuring late projects. But before that price is reached, some banks fail. Bank failures then subtract liquidity from the system, ensuring that when the price is reached, there is too little liquidity for the remaining banks to survive. If instead of many small type H banks, there was a single large type H bank (and aggregate deposits were 0.976 or less), it would be willing to lend at a subsidized rate to the type L banks, to prevent their joint demise. But with many H type banks, self-interest can overcome collective interest and cause a bank meltdown. Note that such a meltdown occurs even in the most optimistic equilibrium.

PROPOSITION 2: *Even if the banking system can generate enough liquidity to pay off date-*0 *investors if all late projects are restructured, the unique (and most optimistic equilibrium) can be one in which the entire banking system melts down.*

Proof: Established by example.

The point is that if all banks restructured their late projects, there could be enough liquidity for all banks to pay claimants. But by the time the interest rate is high enough that all banks find it worthwhile to restructure, some banks can fail. Because they subtract liquidity on net, it may no longer be possible to meet liquidity demand even if all banks restructure. As a result, all banks fail.

F. The Causes of Contagion

It is useful to list the assumptions that lead to the possibility of a contagious meltdown. First, the underlying source of friction in this model is the nontransferability of specific skills. Because it has specific skills in collecting on bank loans, a bank has to issue claims demandable at date $\frac{1}{2}$ to deter it from negotiating down depositors at that date. But the very difficulty of renegotiating these claims turns from virtue to vice when there is an anticipated shortfall in available aggregate goods relative to demand. The illiquidity of the bank's loans makes them hard to sell, so banks are forced to restructure projects. Restructuring of late projects typically creates additional liquidity, so this by itself cannot be the source of the spillover. We need a second assumption: that the information about the aggregate state can arrive earlier (at date $\frac{1}{2}$) than when liquidity is produced (date 1). Now anticipation of a future liquidity shortage (or bank insolvency for any other reason) causes depositors to redeem their claims immediately. This forces failing banks to restructure even early projects. The rigid contract structure can lead to an undesirable scramble for liquidity. Note that the reason banks themselves have the right to restructure projects (call loans) at date $\frac{1}{2}$ is to deter borrowers from seeking concessions from the bank on that date.

Even the restructuring of early projects does not cause contagion if the excess demand does not increase in the marginal failure. In particular,

LEMMA 2: *If* $\gamma = 1$ *and* $c_2 = 0$, *then* $(dED/dr) < 0$.

Proof: Obtained by differentiating the excess demand functions over all the ranges.

Excess demand increases with interest rates only if by failing, a bank absorbs more liquidity than if it is solvent. The two sources through which the failed bank reduces what is left for other banks are as follows. First, early entrepreneurs are restructured, leaving them nothing to reinvest, and second, the failed bank absorbs more liquidity than it produces when it sells the date-2 portion of restructured loans. The first effect is absent when $\gamma = 1$ (so that early entrepreneurs contribute no liquidity outside their repayment to their bank) and the second effect is absent when $c_2 = 0$. Hence Lemma 2 indicates when there will be no spillover effects from a bank's failure.

In sum, the function of the bank is to hold assets that are hard to sell, which forces it to finance with demand deposits. Unfortunately, this means that when there is an anticipated aggregate shortage and the bank is run, there is an unnecessary demand for immediate goods that the bank can best meet by producing liquidity inefficiently. Thus, simple delays in production can be magnified by banks into a much more severe reduction in aggregate production. This "real" consequence of bank failure stems from the illiquidity of bank assets, and would not be present if bank assets were more saleable. By forcing excessive restructuring, bank failures can cause contagion: They subtract liquidity from the system, which could trigger runs at other banks as well. These spillovers could lead to a unique meltdown equilibrium.[12]

[12] Readers may wonder how important the capital requirement k is to our results. Clearly, if $k = 0$, then outside investors do not naturally "discount" late projects. Nevertheless, if there is an aggregate shortage of liquidity, payoffs from late projects are discounted at rate r, and banks with late projects are worth less than early projects. Banks with excessive late projects are run, and contagion can occur. If, however, there is no aggregate shortage of liquidity to begin with, we require late projects to have lower payoffs than early projects in order for banks to fail. This is not an unreasonable assumption and would again give us contagion. A positive capital requirement is, however, important in adding richness to the model by giving us a region where banks do not want to restructure but the market prefers them to. As Diamond and Rajan (2000) indicate, the capital requirement can be justified as the optimal capital structure issued by a bank faced with uncertainty. While we have not modeled that uncertainty at date 1, the analysis in Diamond and Rajan suggests that the other effects of the uncertainty are orthogonal to the issues we explore in our model.

G. Multiple Equilibria

Finally, let us consider equilibria other than the most optimistic one for a plausible shape for the excess demand function. Let i' be such that $\alpha^{i'}\gamma C + (1-\alpha^{i'})c_1 = d_0$. So banks with $i \geq i'$ are self-sufficient in liquidity and never fail. They never need to raise new deposits and can survive arbitrarily high real interest rates demanded by date 1 investors. Further, let $\hat{r}$ be the inflexion point, so that the excess demand for liquidity decreases monotonically for $r < \hat{r}$ and the excess demand increases monotonically with the interest rate for $r > \hat{r}$, until the interest rate is $\tilde{r}$ where the marginal solvent bank is i' (after this, no more banks fail, and the excess demand for liquidity has to fall with increases in the interest rate).

LEMMA 3: *(i) If there is some $r^* \leq \hat{r}$ such that one of* (6), (7), (8), *or* (9) *hold, then one equilibrium interest rate is r^*. If there is an excess supply of liquidity at $\tilde{r}$, then r^* is the unique equilibrium interest rate. If there is an excess demand for liquidity at $\hat{r}$, then there are two more equilibrium interest rates, one between $\hat{r}$ and $\tilde{r}$, and the other above $\tilde{r}$. (ii) If there is no equilibrium rate r^* that is less than $\hat{r}$, then the unique equilibrium interest rate is above $\tilde{r}$, and all banks that are not self-sufficient in liquidity fail. If there are no self-sufficient banks, the banking system melts down completely.*

Proof: Omitted.

Because of contagion, the unique equilibrium may be a meltdown of all nonself-sufficient banks in the system (as we saw in Proposition 2). For other parameter values, multiple equilibria are possible, one of which again involves runs on all banks that are not self-sufficient. In this case, the meltdown is a self-confirming flight to quality. All expect only the self-sufficient banks to survive, and that the date 2 portion of restructured assets will be highly discounted, offering very high rates of return to those who use their spare goods to buy them. These high rates of return are above those that nonself-sufficient banks can offer. Since the self-sufficient banks do not need new deposits, only they will survive given these beliefs, and the rate of return for surviving early entrepreneurs from buying during the fire sale instead of depositing will indeed be very high. An equilibrium where expectations of high real interest rates cause all banks to fail because none are self-sufficient is Pareto-dominated by the most optimistic equilibrium. Intervention here to select the lower rate will make everyone better off. Of course, such an intervention does not work when the unique equilibrium is a meltdown.

H. The Date-0 Decision: Lending Versus Storage

Thus far, we have ignored the storage decision. Since banks have the option of storing and can finance storage entirely with deposits (they do not have to hold capital against storage), would they optimally store so as to avoid runs? It turns out that

PROPOSITION 3: *In equilibrium, banks either store all the funds they raise or store nothing.*

Proof: See the Appendix.

One might think that the bank might want to avoid a destructive run by storing an additional unit of consumption good at date 0 instead of lending it, so that it is available if the bank's α turns out to be low. But an equivalent action is to reduce the level of outstanding deposits by one unit, and instead raise the balance with capital. Given that the bank can use its capital structure to manage the risk of default, it chooses its asset portfolio so as to maximize returns for a chosen capital structure. Since the expected returns from storage or lending are linear for a given capital structure, the bank prefers either lending everything or storing everything, whichever generates more returns.[13]

I. Example 2, Continued

We have not examined how the level of deposits, d_0, is optimally set ex ante. This problem is one of constrained maximization that offers insights largely peripheral to the focus in this paper, which is on what happens if in the realized state, the level of deposits exceeds available liquidity. Obviously any situation that leads to a meltdown with the optimal level of deposits must have a low ex ante probability. We illustrate this with one of the many ex ante distributions of states of nature that make the level of deposits in Example 2 an ex ante optimum.

Suppose two states of nature are anticipated at date 0. In the good state, all projects are early and $\alpha^i = 1$ for all banks. In the bad state, the distribution of bank types is as in Example 2. Let the probability of the good state be 0.974. In the bad state, all banks fail and investors consume $c_1 = 0.4$. In the good state, entrepreneurs repay banks $\gamma C = 1.28$ at date 1, so investors consume $d_0 = 1$ from deposits plus $\frac{1}{2}(1.28 - 1) = 0.14$ from capital, a total of 1.14. The ex ante capital ratio is then $0.974 \times 0.14/[0.974 \times 1.14 + 0.026 \times 0.4] = 0.12$. The ex ante expected repayment from investing in a bank is $0.974\frac{1}{2}[1.28 + 1] + 0.026 \times 0.4 = 1.12$.

The capital requirement does not permit a higher level of initial deposit issue than $d_0 = 1$. A small reduction in initial deposits issued would only reduce the amount received by investors because banks would still melt down in the bad

[13] From the perspective of the banking system, storage by individuals factors out, since individuals consume all that they store. Also, it might seem strange that even anticipating an extremely high rate of return from buying assets in a fire sale during a meltdown, banks do not store. This stems from the fact that investors care only about date-1 consumption, and in the date-0 competition for funds, banks set initial lending, storage, and capital structure decisions to pay out the maximum to date-0 investors. Storage indeed offers a high return, but that return is only available at date 2. So long as the expected returns to lending between date 0 and date 1 are higher, lending dominates storage, no matter what the interest rate between date 1 and date 2. This is clearly a special feature of the model, but relaxing investor preferences only smooths date-0 allocation between storage and investment and does not alter the results qualitatively. Similarly, with some scale effects in the returns to storage or lending, we would get an interior solution. But what we can show is that even though bank runs make the returns to lending nonlinear in the realizations of α, the pledgeable value is not nonlinear in the fraction lent, which is all that is important for the date-0 decision.

state while investors would get less in the good state. But a significant reduction in the level of deposits issued could allow at least the type H banks to survive in the bad state, and an even further reduction would also allow the L banks to survive (at the cost of the banker absorbing more rents). It is easily checked that the amounts the banking system is able to pay out are below what it can pay with $d_0 = 1$, so this is the constrained optimum. When the bad state is rather unlikely, the trade-off between a stronger commitment to pay in the good states by issuing deposits versus greater stability in the bad states by issuing more capital is dominated by the commitment to pay.

Finally, note that we do not allow for deposits to be made contingent on the interest rate or the nonverifiable state of nature. It is easy to construct a number of states in which the same interest rate prevails but the optimal level of deposits is quite different. Given that there is not a one-to-one mapping between the state and the interest rate, and given that banks compete to pay out the maximum possible, we can provide examples where the ex ante equilibrium interest rate contingent deposit structure creates liquidity problems or solvency problems for some banks in some states. This may even result in a positive (but low) ex ante probability of a complete meltdown of the banking system. We also do not allow contracts to be directly contingent on the state, which we assume is observable but not verifiable. Our analysis is positive—to show what happens when there is an ex post solvency problem or an aggregate liquidity shortage given the use of demand deposit contracts. Obviously, if there was no uncertainty about the ex post state of nature, or if there were complete markets, no such conditions would arise.

J. Is the Assumption That Bank Loans Are Repayable on Demand Realistic?

An important assumption in the paper is that bank loans are repayable on demand. While clearly not all loans need to be callable for the results of the model to apply, how realistic is this assumption today when the received wisdom is that bank loan maturities are fairly long? How realistic is it likely to have been in the past?

In the Federal Reserve Board's February 1997 Survey of the Terms of Bank Lending (Board of Governors (1997)), 84.3% of commercial and industrial loans had a maturity of less than 1 year. Of these short-term loans, 28% were repayable on demand, 25% were overnight loans, and 22% were 1 month and under (not including the first two categories). Therefore, approximately 45% (53% of 84%) of commercial and industrial lending is on extremely short term, and a further 18% (22% of 84%) is very short term. These seem sizeable enough for the model to have relevance even today.

Going back in time, the data become patchier. But we have data comparing short-term bank loans to total loans during the Depression. Short-term loans were near 100% in 1932 and 98% in 1933 (Friend (1954)). Unfortunately, despite extensive search, we have not been able to find the proportion of the short-term loans then that were callable. Nevertheless, it is interesting that bank loan maturity was particularly short during the banking crisis of 1933. Certainly, there was awareness that the short maturity of loans had created problems for borrowers:

> At the same time, businesses were vividly aware of the pressure that had been put by commercial banks during the contraction period of 1929–1932 to liquidate their short term obligations. As a rule these credits had been extended with a realization both by banker and business executive that the funds would be used to finance industrial operations of type that would not conveniently permit of complete liquidation at their legal dates of maturity. The runs on banks during the banking crisis compelled bank management to convert asserts into cash at as rapid a rate as possible, even at the cost of severe embarrassment of insolvency to their business lenders (Jacoby and Saulnier (1942, p. 16)).
>
> In short, the great depression following 1929 heightened the weakness latent in the old business credit forms. Term loans developed partly in recognition of the need for rewriting credit terms to conform more closely to the economic nature of the underlying business transactions, thus protecting the economy from the shock of sudden enforcement of unsuitable credit terms (Jacoby and Saulnier (1942, p. 17)).

Perhaps as a result of this recognition, the proportion of short-term loans to total loans fell steadily till 1940, after which it stabilized at about 70% for a few years and then started climbing again in the 1950s (Friend (1954)). This suggests that even though the extremely high proportion of short maturity bank loans in the early 1930s was somewhat reversed by the move to more term lending, as evidenced by the above quote, the inherent nature of bank lending simply did not allow maturities to be extended beyond a point.

III. Intervention at Date 1

The collective action problem inherent in demand deposits obviously leads to ex post inefficiency. But in addition, failing banks do not internalize the adverse spillover effects they impose on the system as they scramble for liquidity. And the interest rate does not reflect the true value of liquidity since failing banks do not have purchasing power. As a result, solvent banks do not see the true value of liquidity and fail to restructure late projects, even while failing banks are restructuring early projects. Because of these problems, a central authority may find it attractive to intervene.

Since there are no spare resources held outside the system, any intervention will necessarily require taxing some groups and transferring to another. We do not allow the central authority to have any special power of extraction. For example, the central authority can tax resources that are already invested in the banking system, say by existing depositors, but cannot force entrepreneurs to produce or to abstain from consumption. So once entrepreneurs have money left over after repaying loans, the central authority has to make it worthwhile for them to redeposit it in the system.

The central authority can increase the effective supply of date-1 liquidity by taxing claimants on liquidity and lending this back to the system at an interest rate lower than the taxpayer would choose to lend voluntarily. If such a loan is at the market rate of interest that prevails given the loan, we define the tax and loan operation as a *pure liquidity infusion*. If instead the central authority uses its taxation power to provide a particular bank with a gift of future value, for example, a claim on date-2 goods, we define the tax and transfer operation as a *pure recapitalization*. All financial market interventions can be viewed as some combination of pure liquidity infusions and recapitalizations. For example, a gift of current goods to a bank is a liquidity infusion equal to the quantity of current goods, plus a recapitalization equal to the date-2 value of those goods, evaluated at the market interest rate that prevails given the gift.

It is best to think of interventions as unanticipated. Even if we ignore the effects of interventions on the ex ante incentives of the banks and investors (or potential ex post bargaining between the bank and the intervening authority), it is by no means clear that the right intervention is easily identified. The proximate cause of a run is that a bank is deemed unable at current interest rates to raise enough to pay off depositors—it is insolvent. But insolvency can both cause and be caused by liquidity problems. As a result, the effects of interventions that attempt to address either the insolvency of particular banks or the liquidity of the banking system need to be evaluated considering the full general equilibrium, otherwise they may have surprising or unintended effects. This is what we show now.

A. Liquidity Infusion

In principle, liquidity can be raised by taxing any date-1 holdings of consumption goods, but since early entrepreneurs and goods-rich banks already infuse all their consumption goods into the system (provided there is no solvency problem), the taxes have to fall on depositors or capital. Taxing the date-1 claimants after they withdraw consumption goods from the bank ensures that the proper tax is collected without destroying bank incentives to pay depositors. The proceeds of this tax may then be redeposited in solvent banks at random or used to buy the date-2 portion of restructured loans sold by insolvent banks (similar to methods used by the RTC to resolve the savings and loan problems of the 1980s in the United States). If the root cause of the banking system's problems is a liquidity shortage, the liquidity infusion reduces the excess demand for liquidity at date 1 and brings down the interest rate, which will make banks that were insolvent at the prior equilibrium rate solvent again, and eliminate runs.

A pure liquidity infusion does not require particular investors or banks to be targeted. But clearly, relative to the equilibrium interest rate that would prevail if there had been no intervention, net borrowers at date 1 benefit while net lenders are hurt. And, of course, the taxpayers who fund the infusion lose out. Liquidity infusions do alter property rights.

A.1. Interest Rate Guidance: Potential Liquidity Infusion

When multiple equilibrium real interest rates are possible, the central authority could coordinate expectations at the lowest equilibrium interest rate by being prepared to lend and borrow at that rate.[14] This transaction is a promised pure liquidity infusion (evaluated at the low equilibrium rate), which need never occur in equilibrium. While the central authority may potentially need significant taxation authority to give it credibility that it can take sufficiently large positions, in equilibrium it does not need to raise resources. Even though no resources are used, unless the alternative is a complete meltdown, intervention again does not make everyone better off. Early entrepreneurs in banks that do not fail, and banks with excess liquidity, are made worse off (because intervention lowers the rate of return earned on their excess date-1 consumption goods).

B. Recapitalization

The government can recapitalize a bank at date 1 by giving it claims on date-2 consumption goods. Recapitalizations have to be targeted since they assure the solvency of specific banks. An infusion of capital into one part of the banking system does not spread through the system to help other banks survive (in fact, as we will see, it can cause the reverse).[15] When there is not a shortage of aggregate liquidity, the effect of recapitalization is only on the banks receiving capital (see Diamond (2001)).

The government can obtain the funds needed for a pure recapitalization by taxing date-2 wealth. For instance, it can tax banks that have a high realization of α^i by writing deposit claims against them, and granting the deposit certificates to banks that have few borrowers with early projects (a low realization of α^i). Alternatively, it can offer bonds maturing at date 2 to insolvent banks, with the payment on the bonds being funded by a tax on maturing claims at date 2.

A recapitalization may save particular banks when a liquidity infusion will not. For example, suppose the interest rate is 1 at date 1, but banks fail because they are insolvent at that rate. If the government wants to save them, a recapitalization allows them to survive and attract liquidity, but an infusion of liquidity into the system is no help. Since there is already enough liquidity and the interest rate cannot go below 1, additional liquidity just goes to waste.

B.1. Recapitalization in Order to Avoid a Meltdown

If the purpose of intervention is to prevent contagion, the systemic benefit from recapitalizing a bank depends on its characteristics. Holding interest rates constant, if enough capital is provided so that a bank with α^i early borrowers can survive when it restructures a fraction μ^i of its late borrowers, the

[14] If the excess demand is falling with interest rates at the first interest rate equilibrium, then it is sufficient for the central bank to lend at any rate below the next higher equilibrium interest rate.

[15] If the market value of the banking sector as a whole, when no late projects are restructured, exceeds the aggregate value of deposits, then a transfer of value (such as a forced merger) can be enough to recapitalize the insolvent banks. If that market value is insufficient, then value must be transferred from some other source, such as taxing depositors or entrepreneurs, present or future.

excess demand for liquidity will decrease by $\alpha^{i^*}[(c_2/r)+(1-\gamma)C]-(1-\alpha^i)[(1-\mu^i)(\gamma C/(1+k)r)+\mu^i\frac{c_2}{r}]$ times the mass of banks of type α^i. Saving banks with low values of α^i does not decrease the excess demand for liquidity, and at low interest rates may in fact increase it (because the banks prefer to set $\mu^i = 0$ and not restructure late borrowers). For a sufficiently high α^i, saving banks of type i decreases the excess demand for liquidity purely by avoiding the liquidity-absorbing effects of a run that forces restructuring of its early projects. This implies that recapitalizations are best targeted at banks that are intrinsically more liquid in the near term than at banks that are illiquid.

An additional benefit of recapitalization is that it allows high interest rates to prevail without causing the failure of high α^i banks. High interest rates, holding bank failure constant, force a solvent bank to restructure late projects.[16] A generalized recapitalization of all insolvent banks can allow high interest rates to prevail without bank failures. Liquidity is then produced more efficiently, through the restructuring of late projects rather than through the restructuring of early projects.

C. When Do Interventions Work?

In sum, when do liquidity infusions or recapitalizations prevent failure? We have

PROPOSITION 4:

(i) If

$$\int_{\underline{i}}^{\bar{i}} \left[\alpha^i\gamma C+(1-\alpha^i)c_1\right]f(i)\,di + \min\left[\int_{\underline{i}}^{\bar{i}}(1-\alpha^i)c_2 f(i)\,di,\ \int_{\underline{i}}^{\bar{i}}\alpha^i(1-\gamma)Cf(i)\,di\right] \geq d_0,$$

then all banks can be made solvent by an appropriate injection of capital.

(ii) A pure liquidity infusion saves all banks if and only if

$$\alpha^i\gamma C+(1-\alpha^i)\frac{\gamma C}{(1+k)} \geq d_0 \quad \forall i > \underline{i}.$$

Proof: Omitted.

Part (i) states that if there is enough liquidity that can be appropriated by the banking system to pay off deposits, then there is some recapitalization scheme that can allow the liquidity to flow to the right places, no matter what the realization of the α^i's.[17] Part (ii) states that an infusion of liquidity is of no

[16] Another way to achieve this objective is to offer to buy loans to late projects on the condition that they will be restructured.

[17] Note that the goods left with early entrepreneurs after repaying loans are counted as part of the liquidity that can be appropriated by the banking system only if there is enough date-2 value to pay it the storage rate of return. Otherwise, the entrepreneur would store rather than invest in the banking system.

use in saving banks when the problem is purely one of solvency (i.e., banks are insolvent even at an interest rate of 1). But if a liquidity shortage causes solvency problems, then a liquidity infusion can save all banks.

D. Interventions in the Example

Let us now go back to Example 2 to see what interventions can work. In the unique meltdown equilibrium, all banks fail, depositors each consume 0.4, and bankers and entrepreneurs consume a total of 0.5. There are interventions that ex post Pareto dominate this outcome.

A small injection of liquidity can save the type H banks by allowing the interest rate to drop below $r = 2.77$ (the level that makes type H banks insolvent), but it cannot prevent the run on type L banks because they are insolvent at any interest rate in excess of $r = 1.4$. Aggregate liquidity supply is 0.688 (with H banks liquidating all late projects), demand is 0.748 (the value of claims on failed L banks and solvent H banks), implying that a liquidity infusion of 0.06 suffices to save the type H banks. The government should tax 0.06 from depositors after they withdraw, and lend it at date 1 to any bank that can repay at an interest rate of $r = 2.77$. The government should collect repayment at date 2, and possibly distribute these as public goods. If there are no deadweight costs of the tax, expected after-tax consumption is 0.688 at date 1, averaged across all depositors and capital holders.

A large liquidity injection (of 0.251) can allow both types of banks to survive. This allows expected after-tax consumption of 0.768. However, some recapitalization is needed to increase date-1 consumption any further (more liquidity can be produced only by restructuring late projects, and that requires interest rates to exceed 1.6 where L banks are insolvent).

Because deposits exceed the maximum available liquidity (at any interest rate), there can be no pure recapitalization that prevents meltdown. However, capital injections can provide better outcomes when combined with injections of liquidity. If a recapitalization is combined with a large liquidity injection that allows the type L bank to survive at $r = 1.6$ (where L banks restructure late projects to survive), then the liquidity supply can be increased to 0.912. At $r = 1.6$, liquidity demand is 1.011 (with both types of bank solvent), so the liquidity infusion must be $1.011–0.912 = 0.099$. In addition, the pledgeable value of type L banks is 0.94, so to survive they must obtain an additional value of 0.06 in date-1 goods (or 0.06×1.6 at date 2). Since fraction 0.6 of banks are type L, the required recapitalization is $0.06 \times 0.6 = 0.036$. The total cost in date 1 goods is then 0.135 (0.099 in liquidity and 0.036 in capital). This is lower than the cost in date-1 goods of 0.251 for a large liquidity injection, and yields a higher after-tax consumption of 0.912.

Alternatively, a larger capital infusion combined with less liquidity allows type L banks to survive even if the interest rate rises to $\bar{R} = 1.95$, when type H banks supply liquidity by restructuring all late projects. The total cost of this intervention in terms of date-1 goods is 0.087 (0.03 in liquidity and 0.057 in recapitalization) and it increases after-tax consumption to 0.976.

We have not compared the costs of intervention, both in terms of whom they impact, how much deadweight cost they entail, and what the ex ante incentive effects are. Quantifying such costs is a task for future work. The point to note is that as Proposition 4 indicates, there may be no feasible capital infusion that can save the system if the problem is a fundamental liquidity shortage, and vice versa. Furthermore, a mix of the two interventions may be more effective than either alone.

E. Can Intervention Be Harmful?

This suggests that the central authority should worry about the form of intervention it chooses so as to minimize costs. But more worrisome for the central authority is that intervention can make matters worse.

For example, suppose liquidity is tight so that gross interest rates are greater than 1 but less than R and that low α banks fail in the absence of intervention. Suppose now that the central authority bails out the low α banks by recapitalizing them to the point where they do not fail (e.g., by guaranteeing deposits and exercising forbearance). We showed that saving low α banks that do not restructure late projects (because $r < R$) increases the excess demand for liquidity—they would release liquidity if they failed. Enabling these banks to pay higher interest rates on deposits increases market interest rates. Now banks with moderate α may fail. But these banks are more likely to subtract liquidity from the system when they fail, and their failure brings down more borrowers with early projects. Thus, recapitalization of illiquid failed banks may destroy many healthier banks. The failure of the latter may lead to a contagion of failures (see Diamond and Rajan (2002) for an example).

Put differently, a small recapitalization of the worst banks may soon trigger the need for a massive recapitalization. If on embarking on this process of potentially cascading recapitalizations, the central authority does not have the political clout to carry it out to the end, it may violate perhaps what should be the most important rule governing interventions: First do no harm!

Moreover, the central authority may just not be able to identify the problem. For example, it cannot simply use the real interest rate to tell a solvency problem apart from a liquidity problem. It may well be that the root cause of a high real interest rate is a failing bank that is subtracting liquidity from the system. Conversely, it may be that the root cause is too little liquidity to start with.

If one were to adopt the principle "First, do no harm!" it would seem from our model that a liquidity infusion is the most benign in terms of its spillover effects. A liquidity infusion cannot trigger a cascade of bank failures, and it may well arrest one. And a liquidity infusion, at least, saves the most solvent banks, without the central authority having to pick and choose. This is consistent with the view expressed in Bagehot (1873), Thornton (1939), Goodfriend and King (1988) that only liquidity assistance should be provided. But as we argue, there are situations when a liquidity infusion may be ineffective. If the central authority can sequence interventions, a recapitalization may then be called for. If it is to be undertaken, our model suggests that a recapitalization of the most illiquid banks can be a mistake, even if these are the ones that are closest to solvency because the liquidity shortage can be exacerbated. Instead, capital should be provided to the most liquid of the failing banks. Finally, the central authority should recognize that an initial recapitalization typically involves further commitments, and should stand ready to provide it.[18]

IV. Conclusion

We have examined how liquidity shortages and solvency problems in banks interact, and how each can cause the other. Interestingly, the possibility of a contagion of banking failures arises precisely because of the very structure of banks—to deal with a commitment problem, they finance illiquid assets with demandable claims. But if deposits cannot be made perfectly state-contingent, this structure can cause or exacerbate a liquidity shortage. When depositor losses are unavoidable, each depositor demands payment. This can force banks to foreclose on loans that otherwise would soon produce real liquidity.

The model suggests that contagious bank failures need not occur only because of panics unrelated to fundamentals of the economy (though we have the possibility of panics in our model). Because bank failures themselves can cause liquidity problems, our analysis suggests that evidence that bank failures can be predicted with economy-wide business conditions is not sufficient to indicate that such failures are benign (see, e.g., Calomiris and Gorton (1991)). Just because they are not sunspots does not mean that bank runs are in the collective interest of depositors in run banks or that they do not have large external effects. In addition, spillovers could work through local deposits and firms. This suggests the possibility of local contagion as well as the possibility of national contagion. Finally, if runs are not instantaneous (e.g., if the state of nature is learned sequentially by different depositors), then the real effects of a bank calling loans (forcing the restructuring of projects) to avoid runs will influence economic conditions that cause subsequent runs on itself or on other banks. Economic conditions before banking crises, although predetermined, are probably not exogenous to the real effects of bank failure.

Our model is a "real" model, with real rather than nominal contracts. We view this as a strength of the model. It helps us focus on the real budget constraint that an economy has to face. To the extent that financial assets derive their liquidity ultimately from liquid real assets, our model is at the appropriate level of abstraction. Countries like Argentina, whose banking system issued

[18] *Suspension of convertibility* is both an injection of liquidity and of capital, financed by the date-0 investors in a bank. The direct effect of suspension is to eliminate a bank's need to pay out immediately. This stops runs if suspension continues until all entrepreneurs can repay their loans. Even if suspension continues only until early entrepreneurs generate goods, it prevents the destruction of liquidity and possible contagion. *Deposit insurance* is akin to a potentially large infusion of capital, where the actual amount of capital provided depends on the interest rate offered by banks. However, when there is excess aggregate demand for liquidity at all interest rates, the banking system fails immediately if the insurance is finite. Without an injection of liquidity, interest rates will spike upwards, but any finite injection of capital does not allow the market to clear.

dollar-denominated deposits, or countries during the gold standard when exchange rates did not adjust as a matter of course, essentially look like the real model we have described. Nevertheless, modern banking systems issue deposits denominated in local currency, and not goods. In Diamond and Rajan (2003), we show that while introducing such financial assets into the framework of this paper adds new insights on the relationship between money, prices, and real activity, it does not alter the basic implication of this paper that banking crises can be caused and propagated by real liquidity shortages.

Finally, we can weaken a number of assumptions without altering this fundamental result. For example, initial investors, some of whom can consume early and others late, or who can substitute between consumption at different dates, can be easily accommodated. More interesting would be to endogenize the propensity to demand liquidity: An economy is particularly fragile if depositors have a tendency to demand liquidity when projects are predominantly late. The converse would be true if liquidity demand were particularly attenuated when projects were delayed. Exploring this is an important extension best left for future work.

Appendix

Proof of Proposition 3: At date 0, the bank chooses the fraction invested in storage, s_0, as well as the quantity of deposits issued, d_0, per unit of resources raised. The bank must pledge out the maximum to investors in order to be able to raise money in the competitive market for funds. Start first by assuming that the bank invests all resources in lending. Also, assume that there is only one aggregate state at date 1. The extension to when there are multiple aggregate states is straightforward.

For every deposit level d_0, there is a date-1 equilibrium interest rate (if there are multiple equilibrium rates, we assume that the lowest one prevails), a marginally solvent bank, $i^*(d_0)$, and equilibrium decisions by solvent banks on how many of their late projects they will restructure, $\mu^i(d_0)$. It is straightforward to integrate across bank types to obtain the amount that is paid out at date 1 to date-0 investors at that level of deposits, d_0. Let this be $v_1(d_0)$.

Let $d_0^* = \arg\max_{d_0} v_1(d_0)\,\text{s.t.}\,\frac{d_0^*}{v_1(d_0^*)} < \frac{1}{1+k}$. This is the value of deposits that maximizes payout subject to the bank's capital meeting the capital constraint. If $v_1(d_0^*) > 1$, then the maximum payout with the banks lending everything exceeds the maximum payout when the banks store everything.

Now it is easy to show that it is optimal for an individual bank to lend everything and to choose a deposit level of d_0^* when all other banks do so. Assume not. Let the bank store $1 > s_0 > 0$. Let the bank issue d^{s} of deposits against the stored assets. Let it issue d^{L} of deposits against the amount lent. Since the return from storage is certain, the spillover effect of storage, after it pays out on deposits issued against it, is as if the bank issued $(d^{\mathrm{L}} - (s - d^{\mathrm{s}}))$ of deposits against its loans (because the certain returns from storage can be fully offset against outstanding deposits, leaving net deposits as all that matters). But we know that the pledgeable amount from lending is maximized when the deposits issued against $(1 - s_0)$ units of loans is $(1 - s_0)\, d_0^*$. Therefore, it must be that

$$d^{\mathrm{L}} - \left(s_0 - d^{\mathrm{s}}\right) = (1 - s_0)\, d_0^*$$
$$\Rightarrow d^{\mathrm{s}} + d^{\mathrm{L}} = s_0 + (1 - s_0)\, d_0^*.$$

But the last equality implies that one can treat the value from storage as going completely to service the deposits issued against storage, and the value from the amount lent as going completely to service deposits and capital issued against the amount lent. Therefore, the pledgeable amount if the bank stores is

$$s_0 + (1 - s_0) v_1\left(d_0^*\right) < v_1\left(d_0^*\right),$$

where the last inequality is because $v_1(d_0^*) > 1$. Thus, the bank can pledge less at date 1 than other banks if it stores, and does not attract funds because only those who wish to consume at date 1 have goods at date 0. It therefore does not store. A similar line of argument explains why the bank does not lend if $v_1(d_0^*) < 1$. Q.E.D.

REFERENCES

Allen, Franklin, and Douglas Gale, 2000, Financial contagion, *Journal of Political Economy* 108, 1–33.

Bagehot, W., 1873, *Lombard Street* (Scribner Press, New York).

Bhattacharya, Sudipto, and Douglas Gale, 1987, Preference shocks, liquidity and central bank policy, in W. Barnett, and Kenneth Singleton, eds.: *New Approaches to Monetary Economics* (Cambridge University Press, Cambridge, MA).

Board of Governors of the Federal Reserve System, 1997, Survey of the terms of bank lending, February.

Caballero, Ricardo J., and Arvind Krishnamurthy, 2001, International and domestic collateral constraints in a model of emerging market crises, *Journal of Monetary Economics* 48, 513–548.

Caballero, Ricardo J., and Arvind Krishnamurthy, 2004, Smoothing sudden stops, *Journal of Economic Theory* (forthcoming).

Calomiris, Charles W., and Gary Gorton, 1991, The origins of banking panics: Models, facts, and bank regulation, in R. Glenn Hubbard, ed.: *Financial Markets and Financial Crises* (University of Chicago Press, Chicago).

Calomiris, Charles W., and Charles M. Kahn, 1991, The role of demandable debt in structuring optimal banking arrangements, *American Economic Review* 81 (3), 497–513.

Diamond, Douglas W., 1997, Liquidity, banks, and markets, *Journal of Political Economy* 105, 928–956.

Diamond, Douglas W., 2001, Should Japanese banks be recapitalized? *Monetary and Economic Studies* 19, No. 2, 1–19 (Bank of Japan). Revised and reprinted as, Should banks be recapitalized? *Economic Quarterly* (Federal Reserve Bank of Richmond) 87 (4), 71–96.

Diamond, Douglas W., and Phillip H. Dybvig, 1983, Bank runs, deposit insurance, and liquidity, *Journal of Political Economy* 91, 401–419.

Diamond, Douglas W., and Raghuram G. Rajan, 2000, A theory of bank capital, *Journal of Finance* 55, 2431–2465.

Diamond, Douglas W., and Raghuram G. Rajan, 2001, Liquidity risk, liquidity creation and financial fragility: A theory of banking, *Journal of Political Economy* 109, 287–327.

Diamond, Douglas W., and Raghuram G. Rajan, 2002, Bank bailouts and aggregate liquidity, *American Economic Review, Papers and Proceedings* 92, 38–41.

Diamond, Douglas W., and Raghuram G. Rajan, 2003, Money in a theory of banking, Working paper, University of Chicago.

Donaldson, Glen, 1992, Costly liquidation, interbank trade, bank runs and panics, *Journal of Financial Intermediation* 2 (1), 59–82.

Friedman, Milton, and Anna Schwartz, 1963, *A Monetary History of the United States 1867–1960* (Princeton University Press, Princeton, NJ).

Friend, Irwin, with the assistance of Vito Natrella, 1954, *Individuals' Saving: Volume and Composition* (Wiley, New York).

Goodfriend, Marvin, and Robert G. King, 1988, Financial deregulation, monetary policy, and central banking, *Economic Review* (Federal Reserve Bank of Richmond) 74, 3–22.

Hart, Oliver, and John A. Moore, 1994, Theory of debt based on the inalienability of human capital, *Quarterly Journal of Economics* 109 (4), 841–879.

Holmström, Bengt, and Jean Tirole, 1997, Financial intermediation, loanable funds, and the real sector, *Quarterly Journal of Economics* 52, 663–692.

Holmström, Bengt, and Jean Tirole, 1998, Private and public supply of liquidity, *Journal of Political Economy* 106, 1–40.

Jacoby, Neil H., and Raymond Saulnier, 1942, *Term Lending to Business* (National Bureau of Economic Research, New York).

Jeanne, Olivier, 2000, Debt maturity and the global financial architecture, Centre for Economic Policy Research, London Discussion paper #2520 (August).

Kiyotaki, Nobuhiro, and John Moore, 2000, Inside money and liquidity, Working paper, London School of Economics.

Smith, B., 1991, Bank panics, suspensions, and geography: Some notes on the contagion of fear in banking, *Economic Inquiry* 29, 230–248.

Thornton, Henry, 1939, An enquiry into the nature and effects of the paper credit of Great Britain (1802), together with his evidence given before the Committees of Secrecy of the two houses of Parliament in the Bank of England (March and April, 1797), some manuscript notes, and his speeches on the Bullion report (May 1811), in F. A. von Hayek, ed.: *The Library of Economics . . . Section 1: Classics of Economic Thought*, 2 (George Allen & Unwin Ltd., London).