Endogenous Financial Fragility and Prudential Regulation

Charles M. Kahn

Department of Finance
University of Illinois
1206 S. Sixth St.
Champaign, IL 61820
Email: c-kahn@uiuc.edu

João A. C. Santos*

Federal Reserve Bank of New York
33 Liberty St.
New York, NY 10045
Tel: (212) 720-5583
Fax: (212) 720-8363
E-mail: joao.santos@ny.frb.org

June 12, 2005

Initial Version: June 2003

*The authors thank seminar participants at the 2003 CFS Conference on Liquidity Concepts and Financial Instabilities, and the Federal Reserve Bank of New York for valuable comments. The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of New York or the Federal Reserve System.

Endogenous Financial Fragility and Prudential Regulation

Abstract

We study the fragility of the banking system and its implications for prudential regulation. In our framework, fragility stems from the interconnections banks establish to protect themselves from liquidity shocks. We show that when banks do not provide payment services they have an incentive to choose the optimal degree of mutual insurance. Under these conditions, the flexibility with which financial assets can be designed and priced causes all market participants to correctly to take into account the economic effects of their own interdependence. When banks provide payment services this flexibility is no longer available. In this case, we show that banks have an incentive to become too interdependent, since some of the beneficiaries of the payments arrangement are unable to compensate the banks for maintaining independence. This creates a justification for a regulatory intervention. We examine possible modes of intervention.

1 Introduction

In a modern economy, financial institutions interact in a bewildering variety of ways. The assets and liabilities of financial institutions include claims to and from other financial institutions, creating a high degree of interdependency. For the most part, these choices are voluntary: no outside authority forces these cross institutional holdings. On the other hand, no outside authority guarantees that these institutions get the degree of interdependence "right."

Some degree of interdependence is inevitable given the role of intermediaries. Intermediaries have the job of using superior information to channel capital from owner to user through creation of financial assets and liabilities. In a world with a high degree of complexity and specialization some intermediaries will have the job of channeling the capital of other intermediaries. Moreover, in a world where the responsibility for effecting payments is largely taken over by private financial intermediaries, the regular course of commercial activity will impose interdependencies on the institutions which take this role.

Do banks have an incentive to set the optimal degree of mutual dependence in liabilities and payments arrangements? If not what should regulators do about it? A recent literature has argued for an inherent fragility in the financial system, and a variety of mechanisms by which fluctuations in the financial system will propagate through the economy. To what extent is this propagation endogenous, and what are its determinants?

With alliterative allure, the term "financial fragility" encapsulates a key strand—indeed the best surviving strand—of the mainstream theoretical economic argument in favor of macroeconomic stabilization policy in a world where prices are flexible. Outside the financial system, the main prescriptions given for increasing economic stability are neoclassical and laissez faire in their outlook: allow the participants in the market to adjust to the changes in the economy by eliminating structural rigidities in labor and product markets; increase transparency so that market participants can correctly and intelligently adjust to government's and firms' activities. In the case of financial markets, however, a respectable body of opinion has always argued, and still argues, for an inherent instability and a need for intervention to correct that instability.

Economists and historians have always looked to the financial system as a possible source of the fluctuations that have hit economies. This focus is natural: the apparent rapid response of financial institutions to news and innovations and the pervasiveness of financial activities in the economy as a whole also makes them a natural source and propagation mechanism for transmitting fluctuations to the economy. Still, even if we accept such claims, we do not know that the financial system delivers the "wrong" degree of fluctuation to the economy. Perhaps the financial system is not "fragile" but "responsive."

In this paper we argue that for the most part, financial institutions have the incentive to get it "right." The flexibility with which financial assets can be designed and priced causes the market participants correctly to take into account the economic effects of their own interdependence. We illustrate this result by showing in a modified version of the Diamond and Rajan (2001) model of banks, that banks choose the correct degree of mutual insurance. However, when financial institutions provide payment services we argue that they have an incentive to become too interdependent. When financial institutions play a role in the payments system the aforementioned flexibility is no longer available. As a result, some of the beneficiaries of the payments arrangement are unable to compensate the banks for maintaining independence. We show this point with our modification of the Diamond and Rajan (2001) model after we include production and a payments system.

The theoretical underpinnings of our model are the "usual suspects" in the information and contracting literature: moral hazard, imperfect information, and externalities. Indeed the first point is simply an illustration of the general principal of constrained optimality in a decentralized ex ante contracting environment.[1] The inefficiency in the second case arises because ex ante contracting is not possible. Nonetheless, we think that the application of this general principle is particularly relevant when financial intermediaries run payments systems. In general financial institutions are able to devise complex and flexible arrangements to cover

[1] See Prescott and Townsend (1984). Allen and Gale (2003a) provide a general presentation in a financial markets framework, and Allen and Gale (2003b) provide a survey of arguments for and against intervention.

numerous contingencies; thus in general, it is natural to model financial contracts in a complete contracting environment. However payments arrangements are intended to be universal. When such systems are decentralized, there is no way for all the potential recipients of payments to agree ahead of time to complex terms for the arrangement. The natural starting point is to assume no ex ante contracting over payments arrangements between pairs of potential transactors.

The rest of the paper is organized as follows. The next section reviews the related literature and compares it with our contribution. Section 3 presents our model ana analyzes the stability of the banking system when banks do not provide payment services. Section 4 extends our model to consider the role of bank payment services. Section 5 discusses the implications of our model for regulation, and section 6 provides a brief conclusion.

2 Related literature

Bryant (1980) and Diamond and Dybvig (1983) were critical to our understanding that a bank is an inherently unstable financial institution. These papers, however, only provide a limited contribution to our understanding of the potential fragility of a banking system. A reason is that in their models, financial instability is associated with bank runs as a self-fulfilling prophecy ("sunspots"). Moreover, since they model the behavior of a single representative bank they are unable to consider the implications of bank interrelationships for industry stability.

Gorton (1985), Postlewaite and Vives (1987), Jacklin and Bhattacharya (1988), Chari and Jagannathan (1988) and Allen and Gale (1998) added to our understanding of the determinants of the stability of a banking system by showing that when there are information frictions the release of information on the bank's financial condition may trigger a run on the bank's deposits. However, as with the previous papers this literature continued to focus on the behavior of a single representative bank. A full-scale financial collapse could now occur, though, if a run on an individual bank provided depositors of other banks with information about aggregate conditions. De Bandt (1995) and Aghion, Bolton and Dewatripont (1999) model two alternative forms of this propagation mechanism. In the model of De Bandt the propagation arises because banks are subject to both an aggregate and an idiosyncratic shock. In the model of Aghion, Bolton and Dewatripont (1999), the propagation occurs because depositors interpret the failure of a bank as a signal of general lack of liquidity in the banking system and withdraw their deposits forcing other banks into bankruptcy.

Rochet and Tirole (1996), Acharya (2000), Freixas, Parigi and Rochet (2000), and Allen and Gale (2000) further expanded our knowledge of the stability of the banking sector by investigating the potential role of the linkages between banks on the stability of the entire industry. The former two papers focus on the role of noncontractual linkages between banks. In the model of Rochet and Tirole (1996) the interdependence between banks arises from peer monitoring. This monitoring is valuable to control bank moral hazard but it may propagate a crisis from one bank to another. If depositors interpret the failure of one bank as a signal that the other banks were not adequately monitored then this may trigger their failure. In the model of Acharya (2000) the linkages arise from banks' incentives to choose correlated portfolios of assets. In his model, the advantage of failing together gives banks an incentive to choose assets which are highly correlated with other banks' assets creating, therefore, the conditions for a joint failure. (The argument is based on posited strategic advantages of failing together).

Freixas, Parigi and Rochet (2000) and Allen and Gale (2000), in turn, provide theories of systemic risk which build on the financial linkages arising from banks' participation in the interbank market. In the model of Freixas, Parigi and Rochet (2000), the financial crises arises as a result of a coordination failure among depositors. They consider a model where banks face liquidity needs because depositors are uncertain about the region where they need to consume. Once depositors find out about their region of consumption, they can either withdraw their endowment and transfer it to that region or have the local bank transfer it for them. Banks execute these orders via credit lines to each other. This minimizes liquidation costs of long-term investments, but makes the banking system unstable. If depositors believe that there will not be enough resources in their region of consumption, their best response is to liquidate their investment in the home location, making it optimal for depositors in other regions to do the same.

In the model of Allen and Gale (2000) the propagation arises because the mutual insurance arrangement that banks establish to protect them against idiosyncratic bank risk does not offer them protection against aggregate bank asset risk. Given that the banks' insurance arrangement makes them mutually dependent, then loss of value in one bank – driven by a liquidity shock for example – can cause sufficient loss of value in a second bank to precipitate a run there, and so on. Key to this result is Allen and Gale's assumption that the aggregate shock is a zero-probability event. In this case the banks are indifferent between a variety of insurance arrangements, that, nonetheless have different implications if the aggregate shock does occur.[2]

Like the last strand of the literature on the stability of the banking system we too focus on the implications of the financial interconnections between banks and abstract from the problem of contagion runs. Our paper is closer to Acharya (2000) in that we are also interested on the determinants of the stability of the banking system. In his model financial fragility arises because banks choose to correlate their portfolios of assets. In our model in contrast, financial fragility arises because of the mutual insurance arrangement that banks establish to protect them from liquidity shocks. We agree that financial institutions do have wide discretion to invest in projects and therefore a temptation to alter the correlation of their projects; nonetheless, fragility through the asset side will likely involve a much slower process than the fragility resulting from interbank liabilities.

In this regard, our paper is also close to Allen and Gale (2000). In contrast to them, however, we drop the assumption that an aggregate shock is a zero probability event. This is important because banks are no longer indifferent about the degree of mutual insurance. To see what happens in this case, suppose that there are a large number of banks. Each bank

[2]Note that if insurance is carried out in a "daisy chain" each bank in the sequence bears the full brunt of the shock, propagating it without dissipating it. If the insurance arrangement were instead universal then shocks would dissipate as losses spread out among the other participants.

has assets whose value can be high or low. If the asset is low and if the bank were on its own a run would occur. As a result, banks find insurance by other banks valuable. Suppose with high probability, the high and low banks cancel each other out and there is no aggregate shock. However, with a small, but non zero, probability there are a few more low banks than high banks. For sufficiently small excess, the optimal arrangement is universal mutual insurance, and the banks will opt for this arrangement. This will have an important implication. All banks fail or stand together. The probability of failure for any one bank is lower than it would be without insurance, but the probability that all banks fail at the *same* time is in fact *higher* than it would be without insurance. The banks themselves, *ceteris paribus,* prefer this universal insurance arrangement. It is less clear that the bank's creditors do so as well. And if banks perform other useful tasks – such as running a payments system – it is possible that the social costs from having all banks fail is more than proportionately greater than the social cost of having some banks fail.

We explore these issues in a framework in which bank failure, following the intuition of Calomiris and Kahn (1991) is expensive but constructive for ex ante incentives. In our environment it is not even necessary to introduce risk aversion as a motive for interbank liabilities: such liabilities can aid in increasing the bank's available capital or reducing the cost of raising capital. This is established in the following section, which begins by modeling banks's decisions independently, and then considers the consequences of banks' decisions to provide cross guarantees. The section concludes by introducing a rudimentary payments system into the framework.

3 The Model

3.1 Independent Banks

All agents are risk neutral; there is no discounting. There are three periods: 0, 1 and 2. We begin with two types of individuals: consumers and bankers. In each period there is a non-storable good ("plain vanilla ice cream") which serves both as investment input and as consumption. Consumers are born at period 0 with an endowment of this good and they are indifferent between consuming at period 1 or 2. Bankers have no endowment of period 0 good, but they are endowed with an exclusive portfolio of investment projects, which transform period 0 good into good in period 1 or 2. We assume there is a large supply of period 0 good, so that the economy-wide required expected return is 0.

If the projects are held by the banker until maturity, they produce a realized value of Y units of period 2 good. However, the portfolio of projects can be liquidated (in a lump-sum fashion) in period 1 for a realized value of X units of period 1 good, where $X < Y$. These values are stochastic as of period 0 and observable, but not verifiable at period 1.

Each bank raises capital by issuing two sorts of liabilities: "deposits" with a total face value of D and "junior bonds" with a total face value of R. For reasons that we will explain next, we assume that if $Y < D$ the bank is liquidated, depositors receive X and the bank and the bond holders receive zero. If $Y > D$ depositors receive D, the bond holders receive $\min\{(X - D)_+, R\}$ and the banker receives $Y - z$ where

$$
\begin{aligned}
z &= D \quad \text{if } D > X \\
&= X \quad \text{if } D + R > X > D \\
&= D + R \quad \text{if } X > D + R
\end{aligned}
$$

This structure for a bank captures the insight of Diamond and Rajan (2001), that demandable debt enables the bank to borrow more than it could borrow under simple debt. It does so by making it more difficult to renegotiate arrangements after the fact, where simple debt will have limited ability to generate funds because of ex post hold up. On the other hand, the difficulty of renegotiation means that the demandable debt arrangement is inflexible in some circumstances, leading to additional expenses in reorganization or liquidation. These two features are reflected in the fact that when the bank is able to pay the depositors but not the junior bond holders, the bond holders are negotiated down to a lower payment without actually liquidating the bank.

The bank's portfolio of projects is drawn from a pool which is unique to the bank. For bank i, the distribution $F_i(Y, X; K_R + K_D)$ of value of the portfolio depends on the total of the amounts of capital raised from deposits K_D and the capital raised from debt K_R. (For convenience, we will assume the distribution is non atomic, and we will suppress the third argument and the subscript when focusing on a single bank). The amount of capital raised through deposits and bonds depends on the terms (D, R) and the implicit return they generate. The return that the depositors will receive is

$$D \int_{Y>D} dF(Y, X) + \int_{Y<D} X \, dF(Y, X).$$

Note that this is independent of the choice of R. The return that the bond holders receive is

$$\int_{D<X<D+R} [X - D] \, dF(Y, X) + R \int_{X>D+R} dF(Y, X).$$

The banker's objective is to choose (D, R, K_D, K_R) so as to maximize

$$\int_{Y>D} [Y - z] \, dF(Y, X)$$

subject to the following two individual rationality constraints for the two classes of investors:[3]

$$D \int_{Y>D} dF(Y, X) + \int_{Y<D} X \, dF(Y, X) \geq K_D \tag{1}$$

$$\int_{D<X<D+R} [X - D] \, dF(Y, X) + R \int_{X>D+R} dF(Y, X) \geq K_R. \tag{2}$$

Issuing deposits imposes an efficiency loss, therefore we wish to raise as much funding as possible through junior bonds. It is clear that the constraints are relaxed by setting R arbitrarily high;[4] thus we can drop R from the problem and constraint (2) simplifies to

$$\int_{D<X} [X - D] \, dF(Y, X) \geq K_R. \tag{3}$$

[3]Note that we are assuming that the choice of investments—that is the choice of the bank's distribution of asset values—is observable at the time of collecting the period zero investment inputs. If not considerations as in Kahn and Winton (2003) apply.

[4]Diamond and Rajan discuss additional considerations which would limit the bond issue. These considerations are not needed for our purposes.

Even so the amount of capital that can be raised through bonds alone is limited to

$$\int X \, dF(Y,X);$$

if additional capital is needed, deposits must be introduced. Henceforward, we will assume that the bank wishes to raise more capital than it can through junior bonds alone—that is, were we to solve the problem assuming $D = 0$, constraint (3) would be binding.

We can also rewrite the banker's objective as follows:

$$\max_{D,K_D,K_R} \int_{Y>D} Y \, dF(Y,X) + \int_{Y<D} X \, dF(Y,X) - K_R - K_D$$

again, subject to (1) and (3). Note that this simplified expression is also the social surplus associated with the bank's activities. Thus an economy of such banks achieves a constrained efficient outcome.

3.2 Cross Guarantees

Banks could in principle engage in a variety of cross-guarantees for reach others' liabilities. The simplest version of this to consider is complete consolidation. Suppose that there are two banks 1 and 2, which are contemplating such a complete set of cross guarantees; effectively, the assets of both banks jointly back the liabilities of each. Clearly such a consolidated bank will have a common face value for bonds and a common face value for deposits. Its objective will be to choose $(D, R, K_D^1, K_D^2, K_R^1, K_R^2)$ so as to maximize

$$\iint_{Y_1+Y_2>D} [Y_1 + Y_2 - z] \, dF_1(Y_1, X_1; K_D^1 + K_R^1) \, dF_2(Y_2, X_2; K_D^2 + K_R^2),$$

where

$$\begin{aligned} z &= D \quad \text{if } D > X_1 + X_2 \\ &= X \quad \text{if } D + R > X_1 + X_2 > D \\ &= D + R \quad \text{if } X_1 + X_2 > D + R \end{aligned}$$

and subject to

$$D \iint_{Y_1+Y_2>D} dF_1(Y_1, X_1; K_D^1 + K_R^1)\, dF_2(Y_2, X_2; K_D^2 + K_R^2) +$$
$$\iint_{Y_1+Y_2<D} [X_1 + X_2]\, dF_1(Y_1, X_1; K_D^1 + K_R^1)\, dF_2(Y_2, X_2; K_D^2 + K_R^2) \geq K_D^1 + K_D^2$$

and

$$\iint_{D<X_1+X_2<D+R} [X_1 + X_2 - D]\, dF_1(Y_1, X_1; K_D^1 + K_R^1)\, dF_2(Y_2, X_2; K_D^2 + K_R^2) +$$
$$R \iint_{X_1+X_2>D+R} dF_1(Y_1, X_1; K_D^1 + K_R^1)\, dF_2(Y_2, X_2; K_D^2 + K_R^2) \geq K_R^1 + K_R^2.$$

As before, the latter constraint simplifies to

$$\iint_{D<X_1+X_2} X_1 + X_2 - D\, dF_1(Y_1, X_1; K_D^1 + K_R^1)\, dF_2(Y_2, X_2; K_D^2 + K_R^2) \geq K_R^1 + K_R^2.$$

and the consolidated bank's objective simplifies to the social surplus:

$$\max_{D,\{K_t^i\}} \iint_{Y_1+Y_2>D} [Y_1 + Y_2]\, dF_1(Y_1, X_1; K_D^1 + K_R^1)\, dF_2(Y_2, X_2; K_D^2 + K_R^2) +$$
$$\iint_{Y_1+Y_2<D} [X_1 + X_2]\, dF_1(Y_1, X_1; K_D^1 + K_R^1)\, dF_2(Y_2, X_2; K_D^2 + K_R^2) - (K_R^1 + K_D^1) - (K_R^2 + K_D^2)$$

As a result we have:

Proposition 1 *Intermediaries will choose the optimal degree of interdependence.*

The proof is immediate: since each bank individually and the banks as a pair will maximize social surplus, it will be profitable for them to provide mutual insurance if and only if such an action increases social surplus. The key feature generating the result is the fact that capital is fairly priced. It is indeed possible that it is less likely that both banks when independent would fail than that the consolidated bank would fail. But the costs of such a failure to lenders are passed on to the bank through the price of debt.

Extensions. We have only discussed the possibility of the two extremes of bank interdependence: complete independence and full mutual insurance. We could imagine a variety of intermediate cases, and with them a variety of limited or contingent recourse between the holders of one institution's liabilities and the assets of the other institution. Kahn and Winton (2003) show in particular that, when the choice of investments is not observable by the debt holders, the introduction of more complex subsidiary structures, with limited recourse among them, can improve a bank's choice of investments. Nonetheless, it will still be the case that a financial institution with these instruments available to it will still pick the constrained socially optimal degree of interdependence between the various affiliated institutions.

The presentation above implicitly incorporates a variety of competitive assumptions: in particular, no bank possesses monopoly power over the capital market. A consolidated set of banks would very likely possess some such market power, and that would be a deterrent to allowing them to join together. In the presence of antitrust regulation on interest rate pricing, it is conceivable that mutual insurance arrangements could provide a back-door to such market power.

In Allen and Gale (2000) a second issue arises: aggregate shocks have the potential to change the shadow value of capital.[5] That paper's "aggregate shock" is in fact concentrated at a single bank. Thus taking the model literally, a solution would be to cut that single bank out of the insurance arrangement in the case of an "aggregate shock." Arrangements in which the degree of insurance provided by one bank to another varied depending on aggregate conditions are extremely plausible—indeed historically documented. Roberds (1995) cite a variety of examples where early clearing houses provided a limited form of mutual insurance which only arose in times of panic.

Nonetheless, it is also plausible that interbank arrangements are imperfect and that some of them would on occasion mistakenly treat aggregate (non-insurable) shocks as if they were insurable. However a systematic restriction of this sort requires quite stringent restrictions on clauses permitted in contracts: First it means treating aggregate shocks as nonbservable, at least within the relevant time frame. It means treating market prices generated—including

[5]See also Diamond and Rajan (2000) and (2003).

pricing of intermediary liabilities themselves—as either unobservable, or as somehow not including aggregate information. And it means prohibiting the use of one firm's bankruptcy as a triggering clause in the contracts of other intermediaries.

3.3 An Example

Suppose that for each of the two banks, Y is distributed uniformly on $[a, b]$ provided that the bank raises at least the amount K in capital. For less capital, $Y = 0$. Furthermore suppose $X = kY$ for a constant k between 0 and 1.

We assume that

$$\frac{a+b}{2} \geq K \geq \left(\frac{a+b}{2}\right) k \tag{4}$$

In other words, the expected payoff from the bank's project exceeds the capital cost, but the funding cannot be raised by bonds alone. In this case, the bank must partly rely on depositors. Define

$$K_D^* = K - \left(\frac{a+b}{2}\right) k$$

the shortfall of capital that could not be raised by bonds alone.

For an independent bank, summing conditions (1) and (3) in the case where F has a uniform distribution establishes that the face value of the deposits must be

$$D = \frac{1}{1-k}\sqrt{2K_D^*(b-a)k}$$

If the value of the project falls below D, then the project goes bankrupt.[6]

Compare the effects if two banks are consolidated. In this case the total amount of capital needed is $2K$, and the same condition 4 determines when issuing deposits is necessary and useful. The consolidated bank's product is now $Y' = Y_1 + Y_2$ where each component has

[6]For this calculation to be correct, we must restrict the parameters so that the resultant D lies in the range $[a, b]$. This and other restrictions on the parameters a, b, k and K are enumerated in the appendix.

an independent uniform distribution. Thus the distribution of Y' is

$$\begin{aligned} F'(Y') &= \frac{1}{2}\left(\frac{Y'-2a}{b-a}\right)^2 \text{ if } 2a \le Y' \le a+b \\ &= 1-\frac{1}{2}\left(\frac{2b-Y'}{b-a}\right)^2 \text{ if } a+b \le Y' \le 2b \end{aligned}$$

A similar calculation to that described above demonstrates that the consolidated firm's debt must have a face value of D' where D' satisfies

$$D'^2\left(D'(1+2k)-6ak\right) = \frac{12K_D^*\,(b-a)^2\,k^2}{(1-k)^2} \tag{5}$$

and provided the solution is in the range $[2a, a+b]$. To determine whether the two firms will consolidate, we must establish whether the expected profits from the consolidated firm are greater or less than twice the expected profits from the individual firm, or equivalently, which of the following is smaller

$$2(1-k)\int_{Y\le D} Y\, dF(Y) = 2(1-k)\int_a^D \frac{Y}{b-a}\, dY = (1-k)\frac{D^2-a^2}{b-a}$$

$$(1-k)\int_{Y'\le D'} Y'\, dF'(Y') = (1-k)\int_{2a}^{D'} Y'\frac{Y'-2a}{(b-a)^2}\, dY' = \frac{1-k}{3\,(b-a)^2}\left(D'^3-3aD'^2+4a^3\right)$$

For the case $a = 0$, the appendix demonstrates that profit is always greater in the consolidated arrangement.

4 Payments and Production

Now we expand our model to include rudimentary production and a payments system. We add a third category of individuals: "producers." There are S different producers. Producers are endowed with a unit of leisure in period 0. Producers only value period 2 (plain vanilla) ice cream and leisure; the opportunity cost of their leisure is c units of period 2 plain vanilla ice cream. By giving up his leisure a producer produces one unit of an idiosyncratic good "specialty flavored ice cream" at period 1. Producers do not meet with other agents until period 1.

We also modify our original model by assuming that depositors (and bond holders) learn at period 1 which producer's output, if any, they wish to consume. A unit of that producer's output yields a utility of $u > c$ (again, measured relative to plain vanilla ice cream); any other producer's period 1 output yields zero utility to that depositor.

Thus a producer wishes to trade specialty ice cream at period 1 for plain vanilla ice cream at period 2. However, ice cream is not attachable: attempts to force depositors to give it up are ineffective. Thus there is no way that ice cream can be used as collateral to back IOU's issued by depositors attempting to buy specialty ice cream.

On the other hand, investments in the bank's projects are attachable; they can be used to back up IOU's (or equivalently "inside money") up to their expected value at any time. Still otherwise put, depositors can use the financial instruments of a solvent bank to pay for idiosyncratic goods.[7]

[7] Note that in this account, both the demand deposits and the junior debt can be used as payment. If we were to refine the model to add asymmetry of information between depositors and producers as to the value of the financial instruments there could be variations in the liquidity of different forms of financial instruments—that is, some financial instruments (e.g. demand deposits) could prove to be superior as means of payment. For theories as to why demandable debt is particularly suitable to payment see Gorton and Pennacchi (1990) and Calomiris and Kahn (1991).

The following table presents a time line of the model.

Period	Actions
0	Deposit contracts are written between banks and depositors Banks make investments Production choices made Banks make mutual insurance arrangements
1	Bank investment results revealed Liquidation decisions made Depositors learn tastes for specialty goods Depositors and suppliers trade
2	Investments mature Second period consumption realized

Suppose that all surplus of the trade is captured by the producer (we will specify an example below where this occurs). From the point of view of the depositors and the bankers, the situation is exactly as described in the previous section, and decisions on debt contracts are unaffected by the production side. However social surplus now includes the expected profits of the producers. Suppose that T consumers have the bank debt enabling them to buy production goods. Then let $\Pi(S,T)$ be the probability that S producers find matches with consuming agents. If all producers produce then the profits to them are

$$\sum_S u\Pi(S,T) - Sc$$

We do not need to know the details of the random process generating Π; all we require is that this process makes the above quantity to increase in T but at a decreasing rate. (Most natural assumptions about random matching processes will satisfy that requirement). When two banks are solvent, there can be $2K/u$ consumers capable of buying production goods. When only one bank is solvent there can be only K/u consumers capable of buying production goods. Note that this means that the social cost of one bank's failure increases when there is a second bank failure. This marginal social cost of interdependence is not included in calculations that banks and debt holders make in determining the profit maximizing degree of interdependence. We conclude:

Proposition 2 *Intermediaries will never choose too little interdependence. For some parameter values intermediaries will choose excessive interdependence.*

In particular, for some parameter values, the choice of excessive interdependence by banks discourages production.

4.1 Example, continued

We return to the uniform-distribution example with two banks. Suppose each individual starts with exactly K units of period 0 ice cream. Thus each bank will have exactly one investor (with investments divided between deposits and junior bonds.) If the bank is solvent in period 1, then each investor has Y in value that can potentially be traded for specialized good. As long as the price of the good is lower than u, the investor is willing to purchase Y/u units of the good. The markets are assumed competitive (meaning that price is determined by market clearing). If the producer produces the good, then ex post he supplies one unit inelastically.

Each individual has a $1/S$ chance of prefering the goods of any particular vendor. All preferences are independent.

Suppose the banks are independent and only one bank is solvent. Then the favored supplier receives payment equal to $\min\{Y, u\}$ from the demander in return for the good. If both banks are solvent, then the supplier receives $\min\{Y_1 + Y_2, u\}$ if the good is demanded by both depositors. Thus the expected payment to the supplier is

$$\frac{1}{S^2}\left[\int_{Y_1>D}\int_{Y_2>D}\min\{Y_1+Y_2,u\}\;dF(Y_1)dF\,(Y_2)+2\int_{Y_1>D}\int_{Y_2<D}\min\{Y_1,u\}dF(Y_1)dF\,(Y_2)\right]$$
$$+2\left(\frac{1}{S}-\frac{1}{S^2}\right)\left[\int_{Y_1>D}\min\{Y_1,u\}dF(Y_1)\right]$$

If the banks are consolidated then the payoff depends on whether the combined bank is solvent:

$$\frac{1}{S^2}\iint_{Y_1+Y_2>D'}\min\{Y_1+Y_2,u\}\;dF(Y_1)dF\,(Y_2)$$
$$+2\left(\frac{1}{S}-\frac{1}{S^2}\right)\iint_{Y_1+Y_2>D'}\min\{\frac{Y_1+Y_2}{2},u\}\;dF(Y_1)dF\,(Y_2)$$

Now the comparison of these two, as well as the decisions by the banks and depositors as to whether to consolidate will be extremely complex in general. But in the case where u is below both D and $\frac{1}{2}D'$ the calculations simplify considerably.[8] In this case the investors get no surplus from the purchase[9] (the supplier gets the entirety of the surplus whenever the bank is solvent) and the above formulas simplify as follows:

$$\begin{aligned}&\frac{u}{S^2}\left[1-(\Pr\{Y_1<D\})^2\right]\\&+2u\left(\frac{1}{S}-\frac{1}{S^2}\right)(1-\Pr\{Y_1<D\})\end{aligned}$$

$$u\left(\frac{2}{S}-\frac{1}{S^2}\right)(1-\Pr\{Y_1+Y_2<D'\})$$

For S large, the crucial comparison is between $\Pr\{Y_1 < D\}$ and $\Pr\{Y_1 + Y_2 < D'\}$, .that is, between $F(D)$ and $F'(D')$. For the uniform distribution on $[0, b]$, the appendix demonstrates that $F(D) > F'(D')$ for large values of k and $F(D) < F'(D')$ for small values of k.

Recall, however, that the banks always prefer to consolidate. The implication is that when k is large, the decision by banks to pool risks increases the probability that the whole banking system fails at once, and thus reduces the probability that the supplier will make a sale. For costs c sufficiently close to u, this means that consolidation will discourage production.

Comments. It is not essential that bank debt be the sole means of payment; it is only necessary that it is the most efficient means of payment.[10] For the point we make it is also not essential to restrict money to demandable debt and therefore we allowed *all* bank debt—demandable or ordinary—to be used for payment. Incorporating such considerations will leave the result unaffected.

[8] Recall that the evaluations of D and D' do not depend on u. Thus we are free to pick parameterizations of u independent of the earlier calculations.

[9] When all banks are insolvent, the market clearing price of the good is zero, but at a price of zero, all consumers in the economy, not just the two holders of shares in the defunct banks, are potential demanders of the specialized good.

[10] For rationales for the existence and efficiency of inside money, see Cavalcanti, Erosa and Temzelides (1999) and Kahn and Roberts (2002).

Note that we have assumed that all market power in the product market rests with the sellers. For this reason no additional contracts between the buyers and their intermediaries will solve the problem. If we were to assume, instead, that all market power rested with the buyers, it would be possible for the intermediaries to extract value from the buyers up front in return for a less fragile arrangement. However in this case, the product market would collapse entirely, since the ex post market price of the good (zero in our example) would not compensate for the up-front costs of its manufacture K. At the expense of notation and obfuscation, the model can be modified to allow for more complicated splits in the case of bilateral bargaining ex post (that is when exactly one buyer is interested in a producer's idiosyncratic good). The results will continue to hold as long as two key features are incorporated: the ex ante market for deposits is competitive, and the ex post price of the good leaves the sellers with *some* surplus. The inefficiency arises from the externality provided by the existence of solvent intermediaries as sources of payments media, and the inability of these intermediary to capture or otherwise to take into account this seller surplus.

5 Implications for Regulation

With only two banks in our example, the inability of sellers to contract with them ex ante strains credulity. As the number of banks in the payment system increases, the idea becomes more plausible. The payments system has a natural scale economy: it is most effective when it is universal, encouraging production by extending the market for product. To the extent that payments systems are decentralized, with producers potentially receiving payment from any of a large number of purchasers delivered through any of a large number of intermediaries, it is unreasonable to assume that a potential recipient of payment can pre contract with all potential customers, or with their representative intermediaries. As a result, there is no incentive for those intermediaries to maintain the effectiveness and stability of that system.

Natural remedies include public maintenance or subsidy for the payments system. A related regulatory remedy is for the government to intervene to protect major paymentsproviding institutions in the case of widespread failure—a policy of "too strategically important to fail." In some cases this subsidy can take the form of ready provision of liquidity in response to aggregate shocks.[11]

In general, the policy implications of this model are similar to those of Acharya (2000): we want to provide banks with incentives not to fail when others are failing—that is, to be more forbearing of idiosyncratic failures than of collective failures. Like Acharya's analysis, ours thus leads to an immediate policy dilemma: the conclusion runs into the face of the standard recommendation that idiosyncratic failure is a signal of malfeasance, and thus needs to be punished more severely. However accounts relying on interdependence of bank liabilities differ from accounts relying on interdependence of bank assets in terms of the operative time scale. The combination of large interbank liabilities, high leverage, and ease of adjusting the financial portions of the balance sheet make the liability side of the balance sheet important for short-term economic stability. Given the relative illiquidity of bank assets the problems of excessive interdependence are therefore more likely to be of relevance for stabilization policy on the liability side.

Our framework emphasizes that it is the *decentralization* of the payments system that causes the potential instability of the financial sector. If the payment system is concentrated in the hands of a small number of institutions—and in particular if the payments system is arranged like the credit card industry, where payments recipients also establish extensive ex ante contractual relations with the payments mechanism (through, e.g., credit card merchant accounts)—then there is scope for the internalization of costs imposed by overdependence among financial institutions. Loosely speaking, then, the arguments we have provided would better justify intervention by financial authorities in America than in Europe.

[11]Ready provision of liquidity by the government is also the remedy in Allen and Gale (2000).

6 Final remarks

Fragility stems from the interconnections banks establish to protect themselves from liquidity shocks. Mutual recourse for bank liabilities is often advantageous. The question we have addressed is whether banks have an incentive to increase fragility beyond the socially desirable level. When aggregate shocks are no longer a zero probability event banks find cross guarantees by other banks valuable, and are no longer indifferent about the form this insurance takes. We have build a framework examining the consequences for interbank insurance when individual bank fragility stems from a moral hazard problem.

As long as ex ante contracting between equally-uninformed parties is feasible, then banks and debt holders can reach constrained optimal arrangements for themselves. If cross guarantees impose a risk, then the price of debt adjusts accordingly. Nonetheless, cross guarantees may be suboptimal, since it places too high a correlation on individual bank failures. If banks also provide payments services, then the social costs may be higher from having all fail than from having some fail, thus creating a justification for a regulatory intervention, when payments systems are decentralized.

7 Appendix

7.1 Consistency of parameter requirements

In the example, the following restrictions must be made on the parameters a, b, K and k.

As noted in the text, the following condition is necessary and sufficient for the firm to need to rase funds from deposits and for the project to be worth funding.

$$\frac{a+b}{2} \geq K \geq \left(\frac{a+b}{2}\right) k \tag{6}$$

In order for D as calculated to be in the interval $[a, b]$ the following condition is necessary and sufficient:

$$a^2 \leq \frac{k\left(2K-(a+b)\,k\right)(b-a)}{(1-k)^2} \leq b^2 \tag{7}$$

We also require D' to be in the interval $(2a, a+b)$, where D' is implicitly defined by equation (5).

We require that the resultant profits in each case be non negative; otherwise the banks will not produce. For the independent bank, this requirement is

$$b^2 - ka^2 - 2K(b-a) \geq \frac{(2K - (a+b)\,k)\,(b-a)k}{1-k}. \tag{8}$$

For the consolidated bank this requirement is redundant provided profits in the consolidated bank exceed those in the independent case.

To show that these conditions are mutually consistent, we consider the case where $a = 0$. Then conditions (6) and (7) can be rewritten as

$$\frac{1}{2} \geq \frac{K}{b} \geq \frac{k}{2}$$

and

$$\frac{1}{2k} - (1-k) \geq \frac{K}{b}$$

It is clear that these conditions are not inconsistent. The condition on D' reduces to

$$\frac{12K^*}{k(1-3k^2+2k^3} \leq b^3;$$

given k, having found a value for K/b satisfying the other conditions, it is only necessary to increase K and b proportionately to find values that satisfy this last condition as well.

7.2 Comparison of profits

For the case $a = 0$, we can establish which profits are larger by evaluating the ratio of the efficiency losses. The efficiency loss of the two banks is

$$(1-k)\frac{D^2}{b} = (1-k)\frac{2kK^*}{(1-k)^2}$$

in the independent case and

$$(1-k)\frac{D'^3}{3b^2} = (1-k)\frac{4k^2K^*}{1-3k^2+2k^3}$$

in the consolidated case.

The ratio is

$$\frac{3bD^2}{D'^3} = \frac{1 - 3k^2 + 2k^3}{2k(1-k)^2}$$

and the numerator exceeds the denominator for $k \in [0, 1]$.

7.3 Probability of bank failures

For $a = 0$

$$F(D) = \frac{D}{b} = \frac{1}{b}\sqrt{\frac{2K^*b}{k + k^{-1} - 2}}$$

and

$$F'(D') = \frac{D'^2}{2b^2} = \frac{1}{2b^2}\sqrt[3]{\frac{2K^*b^2}{(k^{-1} - 1)\left(\frac{1}{2} - \frac{k}{3}\right)}}$$

Taking the sixth power of the ratio of these two shows that independent banks fail less often than banks pooling risk provided that

$$\frac{K}{b} > \frac{2(1 - 3k^2 + 2k^3)^4}{81(1-k)^6k^5}$$

for feasible K/b combinations, this condition is satisfied for large k and violated for small k.

References

Acharya, V.V. (2000), "A Theory of Systemic Risk and Design of Prudential Bank Regulation" Mimeo, Stern School of Business, New York University.

Allen, F. and D. Gale. (2000), "Financial Contagion." *Journal of Political Economy* 108(1), 1-33.

Allen, F. and D. Gale (2003a), "Financial Intermediaries and Markets," Wharton Financial Institutions Working Paper.

Allen, F. and D. Gale (2003b) "Competition and Financial Stability" presented at the World Bank and Federal Reserve Bank of Cleveland project on Bank Concentration.

Bryant, J. (1980), "A Model of Reserves, Bank Runs, and Deposit Insurance," *Journal of Banking and Finance* 4, 335-344.

Calomiris C.W. and C.M. Kahn (1991), "The Role of Demandable Debt in Structuring Optimal Banking Arrangements," *American Economic Review* 81(3), 497-513.

Cavalcanti, R. O., Erosa, A. and T. Temzelides (1999), "Private Money and Reserve Management in a Random-Matching Model," *Journal of Political Economy* 107(5), 929-45.

Chari, V. V. and R. Jagannathan (1988), "Banking Panics, Information, and Rational Expectations Equilibrium," *Journal of Finance* 43, 749-761.

De Bandt, O., (1995) "Competition among Financial Intermediaries and the Risk of Contagious Failures," Notes d'Estudes et de Recherches, No. 30, Paris, Banque de France.

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

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

Diamond, D.W. and R.G. Rajan (2001), "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking," *Journal of Political Economy* 109(2), 287-327.

Diamond, D.W., and R.G. Rajan (2003), "Money in a Theory of Banking," NBER working paper.

Gorton, G. and G. Pennacchi, (1990) "Financial Intermediaries and Liquidity Creation," *Journal of Finance* 45, 49-71.

Jacklin, C. and S. Bhattacharya (1988), “Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications,” *Journal of Political Economy* 96, 568-592.

Kahn, C.M. and W. Roberds (2002) Transferability, Finality and Debt Settlement, Unpublished Manuscript, Federal Reserve Bank Of Atlanta.

Kahn, C.M.. and A. Winton (2003) “Moral Hazard and Optimal Subsidiary Structure for Financial Institutions”, working paper.

Prescott, E. and R. Townsend (1984), “Pareto Optimal and Competitive Equilibria with Adverse Selection and Moral Hazard,” *Econometrica*, 52(1), 21-45.

Rochet, J.C. and J. Tirole (1996), “Interbank Lending and Systemic Risk,” *Journal of Money Credit and Banking* 28(4), 733-762.

Roberds, W. (1995) “Financial Crises and the Payments Systems: Lessons from the National Banking Era,” Federal Reserve Bank of Atlanta *Economic Review* 75, Sept/Oct, 15-31.