Wharton

Financial Institutions Center

Optimal Financial Crises

by
Franklin Allen
Douglas Gale

97-01

The Wharton School
University of Pennsylvania

THE WHARTON FINANCIAL INSTITUTIONS CENTER

The Wharton Financial Institutions Center provides a multi-disciplinary research approach to the problems and opportunities facing the financial services industry in its search for competitive excellence. The Center's research focuses on the issues related to managing risk at the firm level as well as ways to improve productivity and performance.

The Center fosters the development of a community of faculty, visiting scholars and Ph.D. candidates whose research interests complement and support the mission of the Center. The Center works closely with industry executives and practitioners to ensure that its research is informed by the operating realities and competitive demands facing industry participants as they pursue competitive excellence.

Copies of the working papers summarized here are available from the Center. If you would like to learn more about the Center or become a member of our research community, please let us know of your interest.

Anthony M. Santomero
Director

The Working Paper Series is made possible by a generous grant from the Alfred P. Sloan Foundation

Optimal Financial Crises [1]

December 20, 1996

Abstract : Empirical evidence suggests that banking panics are a natural outgrowth of the business cycle. In other words panics are not simply the result of "sunspots" or self-fulfilling prophecies. Panics occur when depositors perceive that the returns on the bank's assets are going to be unusually low. In this paper we develop a simple model of this type of panic. In this setting bank runs can be incentive-efficient: they allow more efficient risk sharing between depositors who withdraw early and those who withdraw late and they allow banks to hold more efficient portfolios. Central bank intervention to eliminate panics can lower the welfare of depositors. However there is a role for the central bank to prevent costly liquidation of real assets by injecting money into the banking system during a panic.

Franklin Allen is at the Wharton School of the University of Pennsylvania, Philadelphia, PA 19104.

Douglas Gale is at the Economics Department of New York University, New York, NY 10003.

I. INTRODUCTION

From the earliest times, banks have been plagued by the problem of *bank runs,* in which many or all of the bank's depositors attempt to withdraw their funds simultaneously. Because banks issue liquid liabilities in the form of deposit contracts, but invest in illiquid assets in the form of loans, they are vulnerable to runs that can lead to closure and liquidation. A financial crisis or *banking panic* occurs when depositors at many or all of the banks in a region or country attempt to withdraw their funds simultaneously.

Prior to the twentieth century banking panics occurred frequently. Panics were generally regarded as a bad thing and the development of central banks to eliminate panics and ensure financial stability has been an important feature of the history of financial systems. It has been a long and involved process. The first central bank, the Bank of Sweden, was established over 300 years ago. The Bank of England played an especially important role in the development of effective stabilization policies in the eighteenth and nineteenth centuries. By the end of the nineteenth century, banking panics had been eliminated in Europe. The last true panic in England was the Overend, Gurney & Company Crisis of 1866.

The U.S. took a different tack. Alexander Hamilton had been impressed by the example of the Bank of England and this led to the setting up of the First Bank of the United States and subsequently the Second Bank of the United States. However, after Andrew Jackson vetoed the renewal of the Second Bank's charter, the U.S. ceased to have a Central Bank in 1836. It also had many crises. Table 1, which is from Gorton (1988), shows the banking crises that occurred repeatedly in the U.S. during the nineteenth and early twentieth centuries. During the crisis of 1907 a French banker commented that the U.S. was a "great financial nuisance". The comment reflects the fact that crises had essentially been eliminated in Europe and it seemed as though the U.S. was suffering gratuitous crises that could have been prevented by the establishment of a central bank.

Eventually the Federal Reserve System was established, in 1914. In the beginning it had a decentralized structure, which meant that even this development was not very effective in eliminating crises. In fact, major banking panics continued to occur until the reforms enacted after the crisis of 1933. At that point, the Federal Reserve was given broader powers and this together with the introduction of deposit insurance finally led to the elimination of periodic banking crises.

Although banking crises may appear to be a thing of the past, it is important to understand why banking panics occurred before central banks devised and implemented policies to prevent them. In an unregulated financial system, banks always have the option of eliminating runs by restricting the ability of depositors to withdraw their funds and by holding sufficient liquid reserves to meet their commitments. Instead, banks found it optimal not to take these measures and allowed bank runs to occur. This raises a number of questions. Why did the banks find it (individually) optimal to allow runs? What ability, if any, does a central bank have that private agents lack and which makes it desirable to intervene to prevent runs? Why was U.S. policy in the last half of the nineteenth century so different from European policy?

The history of regulation of the U.S. and other countries' financial systems seems to be based on the premise that banking crises are bad and should be eliminated. We argue below that there are costs and benefits to having bank runs. The attempt to eliminate runs (or insolvency) completely is an extreme policy that imposes costly constraints on the banking system. In this paper, we try to sort out the costs and benefits of runs and identify the optimal incidence of financial crises.

There are two traditional views of banking panics. One is that they are *random events,* unrelated to changes in the real economy. The classical form of this view suggested that panics were the result of "mob psychology" or "mass hysteria" (see, e.g., Kindleberger (1978)). The modern version, developed by Diamond and Dybvig (1983) and others,[1] is that bank runs are self-fulfilling prophecies. If everyone believes that a banking panic is about to occur, it is optimal for each individual to try to withdraw his funds. Since the bank has insufficient liquid assets to meet all of its commitments, it will have to liquidate some of its assets at a loss. Those who withdraw first will therefore get more than those who wait. Anticipating this, all depositors have an incentive to withdraw immediately. On the other hand, if no one believes a banking panic is about to occur,

[1]See also Bryant (1980) and Waldo (1985).

only those with immediate needs for liquidity will withdraw their funds. Assuming that the bank has sufficient liquid assets to meet these legitimate demands, there will be no panic.

Which of these two equilibria occurs depends on extraneous variables or "sunspots". Although "sunspots" have no effect on the real data of the economy, they affect depositors' beliefs in a way that turns out to be self-fulfilling.[2] According to this view of banking panics, it is optimal for governments to intervene to eliminate panics, using either appropriate central banking policies or deposit insurance. Laisser faire is inefficient because private agents or organizations do not possess the government's power to tax and so cannot prevent the occurrence of equilibria in which bank runs occur. Furthermore, threat of government intervention turns out to be costless, because there are no panics in equilibrium and hence no need for government action.

An alternative to the "sunspot" view is that banking panics are a natural outgrowth of the business cycle. An economic downturn will reduce the value of bank assets, raising the possibility that banks are unable to meet their commitments. If depositors receive information about an impending downturn in the cycle, they will anticipate financial difficulties in the banking sector and try to withdraw their funds. This attempt will precipitate the crisis. According to this interpretation, panics are not random events but a response to unfolding economic circumstances. Mitchell (1941), for example, wrote (p. 74)

> "when prosperity merges into crisis . . . heavy failures are likely to occur, and no one can tell what enterprises will be crippled by them. The one certainty is that the banks holding the paper of bankrupt firms will suffer delay and perhaps a serious loss on collection."

In other words, panics are an integral part of the business cycle.

Gorton (1988) has conducted an empirical study to differentiate between the "sunspot" view and the business-cycle view of banking panics. He finds evidence which is consistent with the view that banking panics are related to the business cycle and which is difficult to reconcile with the notion of panics as "random" events. Table 1 shows the recessions and panics that occurred in the

[2] Postlewaite and Vives (1988) have shown how this can be formally modeled as a unique equilibrium.

U.S. during the National Banking Era. It also shows the corresponding percentage changes in the currency/deposit ratio and the change in aggregate consumption, as proxied by the change in pig iron production during these periods. The five worst recessions, as measured by the change in pig iron production, were accompanied by panics. In all, panics occurred in seven out of the eleven cycles. Using the liabilities of failed businesses as a leading economic indicator, Gorton found that panics were systematic events: whenever this leading economic indicator reached a certain threshold, a panic ensued. The stylized facts uncovered by Gorton thus suggest banking panics are intimately related to the state of the business cycle rather than some extraneous random variable.

In this paper, we have two objectives. The first is to formulate a model that is consistent with the business cycle view of the origins of banking panics, in the same way that the Diamond-Dybvig model of bank runs formalizes the “sunspot” view. Our second objective is to analyze the welfare properties of this model and derive some conclusions for the performance of government intervention. Because banking crises arise in our model from real shocks to asset returns, rather than self-fulfilling prophecies, the welfare conclusions are quite different from the “sunspot” model. Bank runs are an inevitable consequence of the standard deposit contract in a world with aggregate uncertainty about asset returns. Furthermore, they allow the banking system to share these risks in an efficient way. In some circumstances, we can show that an unregulated banking system which is vulnerable to crises can actually achieve the incentive-efficient allocation of risk and investment. In other circumstances, where crises are costly, we show that appropriate government intervention can avoid the unnecessary costs of bank runs while continuing to allow runs to fulfill their risk-sharing function. However, in all of these cases, it is *never* optimal to impose artificial constraints on the banks to prevent runs. This is because, in our model, the only way to eliminate runs is to force banks to hold a large proportion of safe assets, something which is never efficient if banks do not choose to do this on their own.

The model is described in Section II. Our assumptions about technology and preferences are the ones that have become standard in the literature since the appearance of the Diamond and Dybvig (1983) model. Banks have a comparative advantage in investing in an illiquid, long-term, risky asset. At the first date, individuals deposit their funds in the bank to take advantage of this expertise. The time at which they wish to withdraw is determined by their consumption needs. Early consumers withdraw at the second date while late consumers withdraw at the third date. Banks and investors also have access to a liquid, risk-free, short-term asset, represented by a storage technology. The banking sector is perfectly competitive, so banks offer risk-sharing contracts that maximize depositors' ex ante expected utility, subject to a zero-profit constraint.

The main difference with the Diamond-Dybvig model is the assumption that the illiquid, long-term assets held by the banks are risky and perfectly correlated across banks. Uncertainty about asset returns is intended to capture the impact of the business cycle on the value of bank assets. Information about returns becomes available before the returns are realized and when the information is bad it has the power to precipitate a crisis.

We begin our analysis in Section II with a simple case that serves as a benchmark for the rest of the paper. There are assumed to be no costs of early withdrawal, apart from the potential distortions that bank runs may create for risk-sharing and portfolio choice. In this context, we identify the incentive-efficient allocation with an optimal mechanism design problem in which the optimal allocation can be made contingent on a leading economic indicator (i.e., the return on the risky asset), but not on the depositors' types. By contrast, a standard deposit contract cannot be made contingent on the leading indicator. However, depositors can observe the leading indicator and make their withdrawal decision conditional on it. When late-consuming depositors observe that returns will be high, they are content to leave their funds in the bank until the last date. When the returns are going to be low, they attempt to withdraw their funds, causing a bank run. The somewhat surprising result is that the optimal deposit contract is incentive-efficient, that is, it produces the same portfolio and consumption allocation as the optimal mechanism. The possibility of equilibrium bank runs allows banks to hold the incentive-efficient portfolio and produces just the right contingencies to provide incentive-efficient risk sharing.

Banks could eliminate the risk of bank runs by holding large amounts of the safe, short-term asset, but if banks choose not to do so, it is clearly sub-optimal for the government to force them to.

In Section III we introduce a real cost of early withdrawal by assuming that the storage technology available to the banks is strictly more productive than the storage technology available to late consumers who withdraw their deposits in a bank run. A bank run, by forcing the early liquidation of too much of the safe asset, actually reduces the amount of consumption available to depositors. In this case, laisser faire does not achieve the incentive-efficient allocation. However, a simple form of intervention overcomes this problem. Suppose that a bank promises the depositor a fixed nominal amount and that, in the event of a run, the government makes an interest-free loan to the bank. The bank can meet its commitments by paying out cash, thus avoiding premature liquidation of the safe asset. Equilibrium adjustments of the price level at the two dates ensure that early and late consumers end up with the correct amount of consumption at each date and the bank ends up with the money it needs to repay its loan. The incentive-efficient allocation is thus implemented by a combination of a standard deposit contract and bank runs.

Once again, eliminating runs by forcing the banking system to hold large reserves is not desirable. It reduces investment in the risky asset and provides sub-optimal risk sharing

One of the special features of the model is that the risky asset is completely illiquid. Since it is impossible to liquidate the risky asset, it is available to pay the late consumers who do not choose early withdrawal. Section IV introduces an asset market in which the risky asset can be traded. Now the banks may be forced to liquidate their illiquid assets in order to meet their deposit liabilities. However, by selling assets during a run, they force down the price and make the crisis worse. Liquidation is self-defeating, in the sense that it transfers value to speculators in the market, and it involves a deadweight loss. By making transfers in the worst states, it provides depositors with negative insurance. In this case, there is an incentive for the government to intervene to prevent a collapse of asset prices, but again the problem is not runs per se but the unnecessary liquidations they promote.

This model illustrates the role of business cycles in generating bank crises and the costs and the benefits of such crises. However, since it assumes the existence of a representative bank, it must be extended before it can be used to study important phenomena such as *financial fragility* and *contagion.* A discussion of future research and other concluding remarks is contained in Section v.

II. Optimal Risk-Sharing and Bank Runs

In this section we describe a simple model to show how cyclical fluctuations in asset values can produce bank runs. The basic framework is the standard one from Diamond and Dybvig (1983), with two important changes. In our model, asset returns are random and information about future returns becomes available before the returns are realized. As a benchmark, we first consider a case in which bank runs cause no misallocation of assets, because the assets are either totally illiquid or can be liquidated without cost. Under these assumptions, it can be shown that bank runs are optimal in the sense that the unique equilibrium with bank runs supports an incentive-efficient allocation of risk and investment.

Time is divided into three periods $t = 0,1,2$. There are two types of assets, a safe asset and a risky asset, and a consumption good. The safe asset can be thought of as a storage technology, which transforms one unit of the consumption good at date t into one unit of the consumption good at date $t + 1$. The risky asset is represented by a stochastic production technology that transforms one unit of the consumption good at date $t = 0$ into R units of the consumption good at date $t = 2$, where R is a non-negative random variable with a density function $f(R)$. At date 1 depositors observe a signal, which can be thought of as a leading economic indicator. This signal predicts with perfect accuracy the value of R that will be realized at date 2. In Section IIA it will be assumed that consumption can be made contingent on the leading economic indicator, and hence on R. Subsequently, we shall consider what happens when banks are restricted to offering depositors a standard deposit contract, that is, a contract which is not explicitly contingent on the leading economic indicator.

There is a continuum of ex ante identical depositors (consumers) who have an endowment of the consumption good at the first date and none at the second and third dates. Consumers are uncertain about their time preferences. Some will be *early consumers,* who only want to consume at date 1, and some will be *late consumers,* who only want to consume at date 2. At date 0 consumers know the probability of being an early or late consumer, but they do not know which group they belong to. All uncertainty is resolved at date 1 when each consumer learns whether he is an early or late consumer and what the return on the risky asset is going to be. For simplicity, we assume that there are equal numbers of early and late consumers and that each consumer has an equal chance of belonging to each group. Then a typical consumer's utility function can be written

$$U(c_1, c_2) = \begin{cases} u(c_1) & \text{with probability } 1/2 \\ u(c_2) & \text{with probability } 1/2 \end{cases}$$

where c_t denotes consumption at date $t = 1, 2$. The period utility functions $u(\cdot)$ are assumed to be twice continuously differentiable, increasing and strictly concave. A consumer's type is not observable, so late consumers can always imitate early consumers. Therefore, contracts explicitly contingent on this characteristic are not feasible.

The role of banks is to make investments on behalf of consumers. We assume that only banks can distinguish the genuine risky assets from assets that have no value. Any consumer who tries to purchase the risky asset faces an extreme adverse selection problem, so in practice only banks will hold the risky asset. This gives the bank an advantage over consumers in two respects. First, the banks can hold a portfolio consisting of both types of assets, which will typically be preferred to a portfolio consisting of the safe asset alone. Secondly, by pooling the assets of a large number of consumers, the bank can offer insurance to consumers against their uncertain liquidity demands, giving the early consumers some of the benefits of the high-yielding risky asset without subjecting them to the volatility of the asset market.

Free entry into the banking industry forces banks to compete by offering deposit contracts that maximize the expected utility of the consumers. Thus, the behavior of the banking industry can be represented by an optimal risk-sharing problem. In the next three sections we consider a variety of different risk-sharing problems, corresponding to different assumptions about the informational and regulatory environment.

A. The Optimal, Incentive-Compatible, Risk-Sharing Problem

Initially consider the case where banks can write contracts in which the amount that can be withdrawn at each date is contingent on *R*. This provides a benchmark for optimal risk sharing. Since the proportions of early and late consumers are always equal, the only aggregate uncertainty comes from the return to the risky asset *R*. Since the risky asset return is not known until the second date, the portfolio choice is independent of *R*, but the payments to early and late consumers, which occur after *R* is revealed, will depend on it. Let *E* denote the consumers' total endowment of the consumption good at date 0 and let *X* and *L* denote the representative bank's holding of the risky and safe assets, respectively. The deposit contract can be represented by a pair of functions, $c_1(R)$ and $c_2(R)$ which give the consumption of early and late consumers conditional on the return to the risky asset.

The optimal risk-sharing problem can be written as follows:

$$(P1)\begin{cases} \max & & \mathrm{E}[u(c_1(R)) + u(c_2(R))] \\ \text{s.t.} & \text{(i)} & L + X \leq E; \\ & \text{(ii)} & c_1(R) \leq L; \\ & \text{(iii)} & c_1(R) + c_2(R) \leq L + RX; \\ & \text{(iv)} & c_1(R) \leq c_2(R). \end{cases}$$

The first constraint says that the total amount invested must be less than or equal to the amount deposited. There is no loss of generality in assuming that consumers deposit their entire wealth with the bank, since anything they can do the bank can do for them. The second constraint says that the holding of the safe asset must be sufficient to provide for the consumption of the early consumers. The bank may want to hold strictly more than this amount and roll it over to the final period, in order to reduce the uncertainty of the late consumers. The next constraint, together with the preceding one, says that the consumption of the late consumers cannot exceed the total value of the risky asset plus the amount of the safe asset left over after the early consumers are paid off, that is,

$$c_2(R) \leq (L - c_1(R)) + RX.$$

The final constraint is the incentive compatibility constraint. It says that for every value of R, the late consumers must be at least as well off as the early consumers. Since late consumers are paid off at date 2, an early consumer cannot imitate a late consumer. However, a late consumer can imitate an early consumer, obtain $c_1(R)$ at date 1, and use the storage technology to provide himself with $c_1(R)$ units of consumption at date 2. It will be optimal to do this unless $c_1(R) \leq c_2(R)$ for every value of R.

The following assumptions are maintained throughout the paper. The preferences and technology are assumed to satisfy the inequalities

$$\mathrm{E}[R] > 1$$

and

$$u'(0) > \mathrm{E}[u'(RE)R].$$

The first inequality simply states that the risky asset is more productive than the safe asset. This ensures that even a risk averse investor will always hold a positive amount of the risky asset. The second inequality is a little harder to interpret. Suppose the bank invested the entire endowment E in the risky asset for the benefit of the late consumers. The consumption of the early consumers would be zero and the consumption of the late consumers would be RE. Under these conditions, the second inequality states that a slight reduction in X and an equal increase in L would increase the utility of the early consumers more than it reduces the expected utility of the late consumers. So the portfolio $(L, X) = (0, E)$ cannot be an optimum if we are interested in maximizing the expected utility of the average consumer.

An examination of the optimal risk-sharing problem shows us that the incentive constraint can be dispensed with. To see this, suppose that we solve the problem subject to the first three constraints only. A necessary condition for an optimum is that the consumption of the two types be equal, unless the feasibility constraint $c_1(R) \leq L$ is binding, in which case it follows from the first order-conditions that $c_1(R) = L \leq c_2(R)$. Thus, the incentive constraint will always be satisfied if we optimize subject to the first three constraints only.

In fact, the preceding argument shows that the optimal contract satisfies

$$c_1(R) = c_2(R) = \frac{1}{2}(RX + L) \text{ if } L \geq RX$$

and

$$c_1(R) = L, c_2(R) = RX \text{ if } L \leq RX.$$

Thus, we can write the unconstrained optimal risk-sharing problem as follows:

$$\max \int_0^{\bar{R}} 2u\left(\frac{RX+L}{2}\right) f(R)dR + \int_{\bar{R}}^{\infty} (u(L) + u(RX))\, f(R)dR$$

$$\text{s.t. } L + X \leq E,$$

where $\bar{R} \equiv L/X$ is the value of the return on the risky asset at which the liquidity constraint begins to bind. The first-order conditions that must be satisfied by an (interior) optimum are:

$$\int u'(c_1(R)) f(R)dR = \lambda$$

and

$$\int u'(c_2(R)) R f(R)dR = \lambda$$

where λ is the Lagrange multiplier of the constraint $L + X = E$. Under the usual concavity assumptions, these first-order conditions uniquely determine the optimal values of L and X, which in turn determine $\bar{R}$, $c_1(R)$, and $c_2(R)$ through the relationships described above.

Theorem 1 *The solution $(L, X, c_1(\cdot), c_2(\cdot))$ to the optimal risk-sharing problem P1 is uniquely characterized by the following conditions:*

$$c_1(R) = c_2(R) = \frac{1}{2}(RX + L) \text{ if } L \geq RX,$$

$$c_1(R) = L, c_2(R) = RX \text{ if } L \leq RX,$$

$$L + X = E$$

and

$$E[u'(c_1(R))] = E[u'(c_2(R))R].$$

Under the maintained assumptions, the optimal portfolio must satisfy $L > 0$ and $X > 0$.

Proof. See the appendix. ■

The optimal contract is illustrated in Figure l. When the signal at date l indicates that R will be low at date 2 (i.e. $R \leq L/X$) the optimal allocation involves early consumers not consuming very much at date 1 in order to allow consumption to be carried over to date 2 to supplement the low returns on the risky asset for late consumers. When the signal indicates that R will be high at date 2 (i.e. $R > L/X$) then early consumers should consume as much as possible at date 1 since consumption at date 2 will be high in any case. Ideally, the high date-2 output would be shared with the early consumers at date 1, but this is not technologically feasible.

To illustrate the operation of the optimal contract, we adopt the following numerical example.

$$U(c_1, c_2) = \ln(c_1) + \ln(c_2)$$

$$E = 2$$

$$f(R) = \begin{cases} 1/3 & \text{for } 0 \leq R \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

For these parameters, it can readily be shown that (L, X) = (1.19, 0.81) and R = 1.47. The level of expected utility achieved is $\mathrm{E}U(c_1, c_2) = 0.25$

B. Optimal Risk-Sharing through Deposit Contracts with Bank Runs

The optimal risk-sharing problem *(P1)* discussed in the preceding section serves as a benchmark for the risk sharing that can be achieved through the kinds of deposit contracts that are observed in practice. The typical deposit contract is "non-contingent", where the quotation marks are necessitated by the fact that the feasibility constraint may introduce some contingency where none is intended in the original contract. We take a standard deposit contract to be one that promises a fixed amount at each date and pays out all available liquid assets, divided equally among those withdrawing, in the event that the bank does not have enough liquid assets to make the promised payment. Let $\bar{c}$ denote the fixed payment promised to the early consumers. We can ignore the amount promised to the late consumers since they are always paid whatever is available at the last date. Then the standard deposit contract promises the early consumers either $\bar{c}$ or, if that is infeasible, an equal share of the liquid assets L, where it has to be borne in mind that some of the late consumers may want to withdraw early as well. In that case, the early and late consumers will have the same consumption.

With these assumptions, the constrained optimal risk-sharing problem can be written as:

$$P2)\begin{cases} \max & \mathrm{E}[u(c_1(R))+u(c_2(R))] \\ \text{s.t.} & \text{(i)} \quad L+X\leq E; \\ & \text{(ii)} \quad c_1(R)\leq L; \\ & \text{(iii)} \quad c_1(R)+c_2(R)\leq L+RX; \\ & \text{(iv)} \quad c_1(R)\leq c_2(R); \\ & \text{(v)} \quad c_1(R)\leq \bar{c} \text{ and } c_1(R)=c_2(R) \text{ if } c_1(R)<\bar{c}. \end{cases}$$

All we have done, here, is to add to the unconstrained optimal risk-sharing problem *(P1)* the additional constraint that either the early consumers are paid the promised amount $\bar{c}$ or else the early and late consumers must get the same payment (consumption).

Behind this formulation of the problem is an equivalent formulation which makes explicit the equilibrium conditions of the model and the possibility of runs. To clarify the relationship between these two formulations, it will be useful to have some additional notation. Let $c_{21}(R)$ and $c_{22}(R)$ denote the equilibrium consumption of late consumers who withdraw from the bank at dates 1 and 2, respectively, and let $\alpha(R)$ denote the fraction of late consumers who decide to withdraw early, conditional on the risky return R. Since early consumers must withdraw early, we continue to denote their equilibrium consumption by $c_1(R)$.

In the event that the demands of those withdrawing at date 1 cannot be fully met from liquid short term funds, these funds are distributed equally among those withdrawing. Those who leave their funds in the bank receive an equal share of the risky asset's return at date 2.

If a run does not occur, the feasibility conditions are

$$c_1(R)\leq L, c_1(R)+c_{22}(R)\leq L+RX$$

as before. If there is a run, then the early consumers and the early withdrawing late consumers share the liquid assets available at date 1

$$c_1(R) + \alpha(R)c_{21}(R) = L$$

and the late-withdrawing late consumers get the returns to the risky asset at date 2

$$(1 - \alpha(R))c_{22}(R) = RX.$$

Since early consumers and early-withdrawing late consumers are treated the same in a run and all late consumers must have the same utility in equilibrium,

$$c_1(R) = c_{21}(R) = c_{22}(R).$$

If there is no run, then we can assume that $c_{21}(R) = c_{22}(R)$ without loss of generality. These conditions can be summarized by writing

$$\begin{aligned} c_1(R) + \alpha(R)c_2(R) &\leq L \\ c_1(R) + c_2(R) &\leq L + RX \end{aligned}$$

where $c_2(R)$ is understood to be the common value of $c_{21}(R)$ and $c_{22}(R)$.

Our final condition comes from the form of the standard deposit contract. Early withdrawers either get the promised amount $\bar{c}$ or the demands of the early withdrawers (including the early-withdrawing late consumers) exhaust the liquid assets of the bank:

$$c_1(R) \leq \bar{c} \text{ and } c_1(R) < \bar{c} \Rightarrow c_1(R) + \alpha(R)c_2(R) = L.$$

Now suppose that a feasible portfolio (L, X) has been chosen and that the consumption functions $c_1(\cdot)$ and $c_2(\cdot)$ satisfy the constraints of the risk-sharing problem $(P2)$. Then define $\alpha(\cdot)$ by putting 0

$$\alpha(R) = \begin{cases} 0 & \text{if } c_1(R) \leq c_2(R) \\ \frac{L}{c_1(R)} - 1 & \text{otherwise} \end{cases}$$

It is always possible to do so, since feasibility assures us that $c_1(R) \leq L$. Now it is easy to check that all of the equilibrium conditions given above are satisfied. Conversely, suppose the functions $c_1(\cdot), c_{21}(\cdot), c_{22}(\cdot)$ and $\alpha(\cdot)$ satisfy the equilibrium conditions above. There is no loss of generality in assuming that $\bar{c} \leq L$, so $c_1(R) < \bar{c}$ implies that $\alpha(R) > 0$ and it is easy to check that the constraints of the risk-sharing problem *(P2)* are satisfied. This proves that solving the risk-sharing problem *(P2)* is equivalent to choosing an optimal standard deposit contract subject to the equilibrium conditions imposed by the possibility of runs.

When we look carefully at the constrained risk-sharing problem *(P2),* we notice that it looks very similar to the unconstrained risk-sharing problem *(P1)* in the preceding section. In fact, the two are equivalent.

Theorem 2 *Suppose that $\{L, X, c_1(\cdot), c_2(\cdot)\}$ solves the unconstrained optimal risk-sharing problem $(P2)$. Then $\{L, X, c_1(\cdot), c_2(\cdot)\}$ is feasible for the constrained optimal risk-sharing problem $(P2)$. Hence, the expected utility of the solution to (P2) is the same as the expected utility of the solution to (P1) and a banking system subject to runs can achieve incentive efficiency using the standard deposit contract.*

The easiest way to see this is to compare the form of the optimal consumption functions from the two problems. From *(P1)* we get

$$c_1(R) = \min\{\tfrac{1}{2}(L + RX), L\}$$

$$c_2(R) = \max\{\tfrac{1}{2}(L + RX), RX\}$$

and from *(P2)* we get

$$c_1(R) = \min\{\tfrac{1}{2}(L + RX), \bar{c}\}$$

$$c_2(R) = \max\{\tfrac{1}{2}(L + RX), L + RX - \bar{c}\}.$$

The two are identical if we put $\bar{c} = L$. In other words, to achieve the optimum, we minimize the amount of the liquid asset, holding only what is necessary to meet the promised payment for the early consumers, and allow bank runs to achieve the optimal sharing of risk between the early and late consumers. The optimal deposit contract is thus illustrated by Figure 1 with $\bar{c} = L$. The solution to the example introduced above is also unchanged with $\bar{c} = 1.19$.

The total illiquidity of the risky asset plays an important equilibrating role in this version of the model. Because the risky asset cannot be liquidated at date 1, there is always something left to pay the late withdrawers at date 2. For this reason, bank runs are typically partial, that is, they involve only a fraction of the late consumers, unlike the Diamond-Dybvig model in which a bank run involves all the late consumers. As long as there is a positive value of the risky asset $RX > 0$ there must be a positive fraction $1-\alpha(R) > 0$ of late consumers who wait until the last period to withdraw. Otherwise the consumption of the late withdrawers $c_{22}(R) = RX/(1-\alpha(R))$ would be infinite. Assuming that consumption is positive in both periods, an increase in $\alpha(R)$ must raise consumption at date 2 and lower it at date 1. Thus, when a bank run occurs in equilibrium, there will be a unique value of $\alpha(R) < 1$ that equates the consumption of early-withdrawing and late-withdrawing consumers.

C. Standard Deposit Contracts without Runs

We have seen that the incentive-efficient outcome can be achieved by means of a "non-contingent" deposit contract together with bank runs that introduce the optimal degree of contingency. Thus, there is no justification for government intervention to eliminate runs. In fact, if runs occur in equilibrium, a policy that eliminates runs by forcing the banks to hold a safer portfolio must be strictly worse.

It is possible, of course, to conceive of an equilibrium in which banks voluntarily choose to hold such a large amount of the safe asset that runs never occur. Suppose that the incentive-efficient allocation involves no bank runs. Then we know from the characterization of the solution to *(P1)* that $c_1(R) = L$ and $c_2(R) = RX$ for all values of R. If we assume that the greatest lower bound of the support of R is 0, then the incentive-compatibility constraint requires that

$$L = c_1(0) \leq c_2(0) = 0.$$

So the entire endowment is invested in the risky asset, the early consumers receive nothing and the late consumers receive RE. But this means that $X = E$ must maximize

$$u(E - X) + \mathrm{E}[u(RX)]$$

and the first-order condition for this is

$$u'(0) \leq \mathrm{E}[u'(RE)R],$$

contradicting one of our maintained assumptions. Hence, runs cannot be avoided in the optimal risk-sharing scheme.

If the central bank were to prohibit holding portfolios that were vulnerable to runs, this would force the banks to guarantee a constant consumption level $c_1(R) = \bar{c}$ to early consumers, which they can only do by lowering the early consumers' consumption and/or by holding excess amounts of the safe asset. By the earlier argument, when $R = 0$ we have $2\bar{c} \leq c_1(0) + c_2(0) \leq L$ so either $\bar{c} = 0$ or $L > \bar{c}$, neither of which is consistent with the optimum.

Theorem 3 Assuming *that the support of R contains* 0, *the incentive-efficient allocation must allow runs. Hence, an equilibrium in which runs are prevented by central bank regulation is strictly worse than the incentive-efficient allocation.*

Theorem 3 shows that preventing financial crises by forcing banks to hold excessive reserves can be suboptimal. The optimal allocation requires early consumers to bear some of the risk. Figure 2 shows the constrained-optimal contract when the bank is required to prevent runs by restricting its promised payout $\bar{c}$ and increasing the level of reserves L. For the parameter values in our example, it can readily be shown that the constrained-optimal portfolio satisfies $(L, X) = (1.63, 0.37)$ and that $\bar{c} = 0.82$. The level of expected utility achieved is $\mathrm{E}[U(c_1, c_2)] = 0.08$. In comparison with the case where the optimal allocation is implemented by runs, the consumption provided to early consumers is lower except when the return to the risky asset is very low $(R \leq 0.56)$. As a result of this misallocation of consumption between early and late consumers, the ex ante welfare of all consumers is lower than in the second best.

The conclusion of Theorem 3 is consistent with the observation that, prior to government intervention, banks chose not to eliminate the possibility of runs, although it would have been feasible for them to do so. Under the conditions of Theorem 3, any government intervention to curb bank runs must make depositors strictly worse off and, in any case, it cannot improve upon the situation, which is already incentive-efficient according to Theorem 2

D. Unequal Probabilities of Early and Late Consumption

The analysis so far has assumed that the probability of being an early consumer is 1/2. This is a matter of convenience only and it can be shown that with appropriate minor modifications the results above all remain valid when the probabilities of being an early or late consumer differ. To see this suppose depositors are early consumers with probability γ and late consumers with probability $1-\gamma$. The probability of being an early (resp. late) consumer is equal to the proportion of early (resp. late) consumers, so the consumption of each type must be multiplied by γ (resp. $1-\gamma$) in the feasibility constraints. Then the optimal, incentive-compatible, risk-sharing allocation solves the following problem:

$$\begin{array}{lll} \max & & E[\gamma u(c_1(R)) + (1-\gamma)u(c_2(R))] \\ \text{s.t.} & \text{(i)} & L + X \leq E; \\ & \text{(ii)} & \gamma c_1(R) \leq L; \\ & \text{(iii)} & \gamma c_1(R) + (1-\gamma)c_2(R) \leq L + RX; \\ & \text{(iv)} & c_1(R) \leq c_2(R). \end{array}$$

Since γ and $1-\gamma$ appear symmetrically in the objective function and the constraints, they drop out of the Kuhn-Tucker, first-order conditions. The characterization of the second-best allocation follows an exactly similar argument to the one given earlier. The total measure of consumers is now one rather than two, so the optimal consumption allocation is

$$c_1(R) = c_2(R) = L + RX \text{ if } L \geq RX$$

and

$$c_1(R) = \frac{L}{\gamma}, c_2(R) = \frac{RX}{1-\gamma} \text{ if } L \leq RX.$$

With appropriate modifications, all the other arguments above remain valid. Similar extensions are available for the results in the following sections, but for convenience we continue to deal explicitly only with the case $\gamma = 1-\gamma = 1/2$.

III. Costly Financial Crises

A crucial assumption for the analysis of the preceding section is that bank runs do not reduce the returns to the assets. The long-term asset cannot be liquidated, so its return is unaffected. By assumption the safe asset liquidated at date 1 yields the same return whether it is being held by the early-withdrawing late consumers or by the bank. For this reason, bank runs make allocations contingent on R without diminishing asset returns. If liquidating the safe asset at date 1 involved a cost, on the other hand, there would be a trade-off between optimal risk sharing and the return realized on the bank's portfolio.

To illustrate the consequences of liquidation costs, in this section we study a variant of the earlier model in which the return on storage by early-withdrawing late consumers is lower than the return obtained by the bank. Since there is now a cost attached to making the consumption allocation contingent on the return to the risky asset, incentive-efficient risk sharing is not attainable in an equilibrium with bank runs. Government intervention is needed to achieve the second best.

A. Optimal Risk Sharing with Costly Liquidation

Let $r > 1$ denote the return on the safe asset between dates 1 and 2. We continue to assume that the return on the safe asset between dates 0 and 1 is one. This assumption is immaterial since all of the safe asset is held by the bank at date 0. As before, one unit of consumption stored by individuals at date 1 produces 1 unit of consumption at date 2. It will be assumed that the safe asset is less productive on average than the risky asset, that is,

$$\mathrm{E}[R] > r.$$

The characterization of the incentive-efficient deposit contract follows the same lines as before. The bank chooses a portfolio of investments (L, X) and offers the early (resp. late) consumers a consumption level $c_1(R)$ (resp. $c_2(R)$), conditional on the return on the risky asset. The deposit contract is chosen to maximize the ex ante expected utility of the typical consumer. Formally, the optimal risk-sharing problem can be written as:

$$(P3)\begin{cases} \max & \mathrm{E}[u(c_1(R))+u(c_2(R))] \\ \text{s.t.} \quad \text{(i)} & L+X\leq E; \\ \quad\quad \text{(ii)} & c_1(R)\leq L; \\ \quad\quad \text{(iii)} & c_2(R)\leq r(L-c_1(R))+RX; \\ \quad\quad \text{(iv)} & c_1(R)\leq c_2(R). \end{cases}$$

The only difference between this optimization problem and the original problem *(P1)* occurs in constraint (iii), which reduces to the earlier formulation if we put $r = 1$.

To solve problem *(P3),* we adopt the same device as before: remove the incentive-compatibility constraint (iv) and solve the relaxed problem. Then note that the first-order conditions for the relaxed problem require

$$u'(c_1(R))\geq ru'(c_2(R)),$$

with equality holding if $c_1(R)<L$. Then $c_1(R)\leq c_2(R)$ for every R, so the incentive-compatibility condition is automatically satisfied.

The arguments used to analyze *(P1)* provide a similar characterization here. There exists a critical value of $\bar{R}$ such that $c_1(R)<L$ if and only if $R<\bar{R}$. Then the consumption allocation is uniquely determined, given the portfolio *(L, X),* by the relations

$$\begin{aligned} u'(c_1(R)) &= ru'(c_2(R)) \text{ if } R<\bar{R}, \\ c_1(R) &= L, c_2(R)=RX \text{ if } R\geq\bar{R}, \end{aligned}$$

where $\bar{R}$ can be chosen to satisfy $u'(L)=ru'(RX)$. With this consumption allocation, we can show, using the maintained assumptions, that the portfolio will have to satisfy $L>0$ and $X>0$ and the first-order condition

$$\mathrm{E}[u'(c_1(R))]=\mathrm{E}[u'(c_2(R))R]$$

together with the budget constraint $L + X = E$ will determine the optimal portfolio.

In the case of the numerical example, it can be shown that if r = 1.05, *(L, X)* = (1.36, 0.64) and $\bar{R}=2.34$. The level of expected utility achieved is $\mathrm{E}[U(c_1,c_2)]=0.32$. Figure 3 illustrates the form of the optimal contract.

B. Standard Deposit Contracts with Costly Liquidation

The next step is to characterize an equilibrium in which the bank is restricted to use a standard deposit contract and, as a result, bank runs become a possibility. The change in the assumption about the rate of return on the safe asset appears innocuous but it means that we must be much more careful about specifying the equilibrium. Let $\bar{c}$ denote the payment promised by the bank to anyone withdrawing at date 1 and let $c_1(R)$ and $c_2(R)$ denote the equilibrium consumption levels of early and late consumers, respectively, conditional on the return to the risky asset. Finally, let $0 \leq \alpha(R) \leq 1$ denote the fraction of late consumers who choose to "run", i.e., to withdraw from the bank at date 1.

The bank chooses a portfolio (L, X), the pair of consumption functions $c_1(R)$ and $c_2(R)$, the deposit parameter $\bar{c}$ and the withdrawal function $\alpha(R)$ to maximize the expected utility of the typical depositor, subject to the following equilibrium conditions. First, the bank's choices must be feasible and this means that

$$\begin{aligned} L + X &\leq E \\ c_1(R) + \alpha(R)c_2(R) &\leq L \\ (1-\alpha)c_2(R) &\leq r(L - c_1(R) - \alpha(R)c_2(R)) + RX. \end{aligned}$$

The first two constraints are familiar. The final constraint says that withdrawals in the last period, which equal the consumption of the late-withdrawing fraction of the late consumers, cannot exceed the sum of the returns on the risky asset and the returns on the part of the safe asset that is carried over to the last period. The reason that we need to take explicit account here of the fraction $\alpha(R)$ of late consumers who withdraw early is that their decision affects the total amount of consumption available. A unit of consumption withdrawn at date 1 reduces consumption at date 2 by $r > 1$, so it is not a matter of indifference as it was under the previous assumption that $r = 1$.

The standard deposit contract requires the bank to pay the depositors who withdraw in the middle period either a fixed amount $\bar{c}$ or as much as it can from liquid assets. Formally, this amounts to saying that

$$\begin{aligned} c_1(R) &\leq \bar{c} \\ c_1(R) + \alpha(R)c_2(R) &= L \text{ if } c_1(R) < \bar{c}. \end{aligned}$$

Finally, we have the incentive-compatibility condition:

$$c_1(R) \leq c_2(R)$$

and the equal-treatment condition:

$$c_1(R) = c_2(R) \text{ if } \alpha(R) > 0.$$

In other words, if some late consumers withdraw in the middle period, their consumption must be the same as the early consumers since they get the same payment from the bank and store it until the last period. In writing down these conditions, we have implicitly assumed that late consumers get the same consumption whether they withdraw early or late. This will be true in equilibrium, of course.

Having specified the constraints, the bank's problem is formally

$$(P4)\left\{\begin{array}{lll} \max & & \mathrm{E}[u(c_1(R)) + u(c_2(R))] \\ \text{s.t.} & \text{(i)} & L + X \leq E \\ & \text{(ii)} & c_1(R) + \alpha(R)c_2(R) \leq L \\ & \text{(iii)} & (1-\alpha)c_2(R) \leq r(L - c_1(R) - \alpha(R)c_2(R)) + RX \\ & \text{(iv)} & c_1(R) \leq \bar{c} \\ & \text{(v)} & c_1(R) + \alpha(R)c_2(R) = L \text{ if } c_1(R) < \bar{c} \\ & \text{(vi)} & c_1(R) \leq c_2(R) \\ & \text{(vii)} & c_1(R) = c_2(R) \text{ if } \alpha(R) > 0. \end{array}\right.$$

In principle, we could solve the bank's problem directly, but it will be convenient to simplify it first.

The simplification requires us to note that the bank is implicitly allowed to choose the equilibrium that will result at dates 1 and 2 and this ensures that runs will not occur unnecessarily.

More precisely,

$$\alpha(R) > 0 \text{ implies that } c_1(R) < \bar{c}.$$

To see this suppose, contrary to what is to be proved, that $\alpha(R) > 0$ and $c_1(R) = \bar{c}$. Now consider an alternative choice for the bank in which $c_2(R)$ is replaced by $\hat{c}_2(R)$ and $\alpha(R)$ is replaced by $\hat{\alpha}(R)$. Put $\hat{\alpha}(R) = 0$ and use constraint (iii) to define $\hat{c}_2(R)$:

$$\hat{c}_2(R) = r(L - \bar{c}) + RX > c_2(R).$$

Since all of the other conditions are satisfied, the fact that $\hat{c}_2(R) > c_2(R)$ contradicts the optimality of $c_2(R)$ and establishes the desired result.

Under the assumption that the bank can select the equilibrium in which no runs occur, if such an equilibrium exists, there are only two cases to be considered. Either there are no runs, $\alpha(R) = 0$, $c_1(R) = \bar{c}$ and $c_2(R) = r(L - \bar{c}) + RX \geq \bar{c}$; or else there are runs, $\alpha(R) > 0$, and $c_1(R) = c_2(R) < \bar{c}$. When there are runs, we have

$$(1 + \alpha(R))c_1(R) = L$$

from constraint (ii) and

$$(1 - \alpha(R))c_2(R) = RX$$

from constraint (iii), so using the equality of $c_1(R)$ and $c_2(R)$ gives us

$$\frac{L}{(1 + \alpha(R))} = \frac{RX}{(1 - \alpha(R))}$$

or

$$\alpha(R) = \frac{L - RX}{L + RX}.$$

Substituting this value into the expression for $c_1(R)$ yields

$$c_1(R) = \frac{L}{(1 + \alpha(R))} = \frac{L + RX}{2}.$$

This is the same expression as we obtained in the costless case, which is not surprising once we recall that none of the safe asset is being held by the bank between dates 1 and 2 when there are bank runs.

Since we know that runs occur if and only if $c_1(R) < \bar{c}$, we know that runs occur if and only if $R < R^*$, where R^* is defined implicitly by the condition

$$\bar{c} = r(L - \bar{c}) + R^* X.$$

In other words, if there are no runs and the early consumers are paid the promised amount $\bar{c}$, there will be just enough to provide the late consumers with a level of consumption that satisfies the incentive-compatibility constraint. Clearly if $R < R^*$ there must be a run because it is not feasible to pay the late consumers $\bar{c}$ and the early consumers cannot get less unless there is a run. On the other hand, if $R \geq R^*$, then it is always feasible to avoid a run and we have shown that in such cases the bank will find it optimal to do so. We focus on the interior case where $R^* > 0$.

Thus, the bank's decision problem can be simplified to the following:

$$\begin{aligned} \max \quad & \int_0^{R^*} 2u\left(\tfrac{L+RX}{2}\right) f(R)dR + \int_{R^*}^{\infty} \{u(\bar{c}) + u(r(L-\bar{c}) + RX)\} f(R)dR \\ \text{s.t.} \quad & \text{(i)} \quad L + X \leq E; \\ & \text{(ii)} \quad R^* = \tfrac{(1+r)\bar{c} - rL}{X}. \end{aligned}$$

There are two types of solution for this problem. The first possibility is $\bar{c} = L$, in which case the optimal deposit contract is the same as the solution to *(P2)* which is illustrated in Figure 1. The second possibility is $\bar{c} < L$. This is illustrated in Figure 4. Note that since $\bar{c} > \frac{L+R^*X}{2}$, the functions $c_1(R)$ and $c_2(R)$ are discontinuous at $R = R^*$. Whether $\bar{c} = L$ or $\bar{c} < L$ it is clear that the first-order conditions for the solution of the incentive-efficient allocation are not satisfied, e.g., for $R < R^*$ the first-order condition $u'(c_1(R)) = ru'(c_2(R))$ is violated. This can be seen directly by comparing Figures 1 and 4 with Figure 3.

The different types of equilibria can be illustrated in the context of the numerical example. As long as $r < 1.25$ the optimal deposit contract is the same as when $r = 1$ because $\bar{c} = L$ and so nothing is invested at rate r between dates 1 and 2. In other words it has $\bar{c} = L = 1.19$, $X = 0.81, \bar{R} = 1.47$ and $\mathrm{E}[U(c_1, c_2)] = 0.25$. For $r \geq 1.25$ the optimal contract has $R^* = 0$. The representative bank finds it optimal to voluntarily prevent runs and the deposit contract is similar to the one shown in Figure 2. For example, in the case of $r = 1.25$, $\bar{c} = 1, L = 1.8$, $X = 0.2$ and $\mathrm{E}[U(c_1, c_2)] = 0.25$. Hence small changes in r around $r = 1.25$ can cause large changes in the bank's optimal portfolio. The final possibility where $\bar{c} < L$ is illustrated by the case where the probability density function of R is uniform on [0, 2.28] rather than [0,3] but everything else is as before. Here for $r = 1.04$, $\bar{c} = 1.30, L = 1.38$, $X = 0.62, R^* = 1.96$ and $\mathrm{E}[U(c_1, c_2)] = 0.045$.

C. Multiple Equilibria

As was noted earlier, the preceding analysis is based on the assumption that, when there are multiple equilibria at date 1, the bank is allowed to select the one that is preferred by depositors. In practice, this means that runs occur only if they are unavoidable, i.e., only if there does not exist an equilibrium without runs. For any portfolio (L, X) and payout $\bar{c}$, it is clear that a run must occur if

$$\bar{c} > r(L - \bar{c}) + RX$$

since it is impossible to pay the early consumers the promised amount $\bar{c}$ and give at least as much to the late consumers. On the other hand, if

$$\bar{c} < r(L - \bar{c}) + RX$$

it is possible to give the late consumers more than $\bar{c}$, so there is an equilibrium without runs. However, if $\bar{c}$ is close enough to $r(L - \bar{c}) + RX$, there is another possibility. If some late consumers decide to run, it will not be possible to pay out $\bar{c}$ at date 1 to the early withdrawers, even if all the liquid asset is paid out, and because the higher return on the safe asset held by the bank is lost through early liquidation, the late-withdrawing consumers will be worse off too. For an appropriate size of run the late consumers will be indifferent between running and waiting.

Let R^{**} denote the critical value of R below which this second type of equilibrium appears. Then R^{**} *is* determined by the condition that

$$\frac{R^{**}X + L}{2} = \bar{c}.$$

If a run occurred at this value of R then it would just be possible to give both types of consumer $\bar{c}$. The fraction of late consumers who run is determined by the condition that

$$(1 + \alpha(R^{**}))\bar{c} = L.$$

A simple calculation shows that

$$(1-\alpha(R^{**}))\bar{c} = 2\bar{c} - L = R^{**}X,$$

so it is just feasible to give the late consumers $\bar{c}$ at date 2. Figure 4 illustrates where R^{**} lies.

For values of R between R^* and R^{**}, we can choose $\alpha(R)$ so that

$$c_1(R) = \frac{L}{1+\alpha(R)} = \frac{RX}{1-\alpha(R)} = c_2(R),$$

which again satisfies the equilibrium conditions and allows a run. Both types of consumers are worse off in this situation than if a run had not occurred, since $c_1(R) = c_2(R) < \bar{c}$, but it is an equilibrium for the given values of L,X and $\bar{c}$ and so it cannot be ruled out. In the context of the numerical example where the probability density function of R is uniform on [0, 2.28], $R^{**} = 1.97$ so there are multiple equilibria for $R \in [R^*, R^{**}] = [1.96, 1.97]$. Of course, if the bank anticipated that this equilibrium would be chosen for some or all values of $R \in [R^*, R^{**}]$, it would choose a different portfolio (L, X) but that would not eliminate the potential multiplicity of equilibria; it would only change the range of values of R for which multiple equilibria exist.

D. Optimal Monetary Policy

The inefficiency of equilibrium with bank runs arises from the fact that liquidating the safe asset at date 1 and storing the proceeds until date 2 is less productive than reinvesting them in safe assets held by the bank. A simple monetary intervention by the central bank can remedy this inefficiency. Essentially, it consists of giving the depositors money instead of goods. In the event of a run at date 1, the central bank gives the representative bank a loan of M units of money. The bank gives depositors a combination of money and consumption whose value equals the fixed amount promised in the deposit contract. Since early consumers want to consume their entire wealth at date 1, they exchange the money for consumption with the early-withdrawing late consumers. The price level adjusts so that the early consumers end up with the second-best consumption level and the early-withdrawing late consumers end up holding all the money. At date 2, the representative bank has to repay its loan to the central bank. For simplicity we assume that the loan bears zero interest. The money now held by late consumers is just enough to allow the bank to repay its loan and the bank has just enough consumption from its remaining investment in the safe asset to give the early-withdrawing late consumers the second-best consumption level. The price level at date 2 adjusts so that the bank and the early withdrawers can exchange money for consumption in the correct ratio and the bank ends up with the amount of money it needs to repay the loan and the consumers end up with the second-best consumption level.

In order for this intervention to have the required effect on the choice of portfolio and the allocation of consumption, the deposit contract has to be specified in nominal terms. This means that a depositor is promised the equivalent of a fixed amount of money D if he withdraws in the middle period and whatever the representative bank can afford to pay in the final period. This intervention does not require the central bank to condition its policy on the return to the risky asset R. It is sufficient for the central bank to give the representative bank an interest-free line of credit which the representative bank can choose to draw on. Whatever part of the line of credit is used must be repaid in the last period. Without loss of generality, we can fix the size of the line of credit from the central bank and assume that the representative bank uses either none or all of it at date 1.

Let (L,X) be the portfolio and let $c_1(R)$ and $c_2(R)$ be the consumption functions derived from the optimal risk-sharing problem *(P3)*. Let D be the nominal value of a deposit at date 1 and let M be the size of the loan available to the representative bank. (We assume that the bank will make use of the full line of credit or none of it). In states in which the consumption of the early consumers is L there is nothing that the representative bank needs to do to prevent runs. As before, in states where $c_1(R) < L$, bank runs are valuable because they make the value of the deposits contingent on R, but here they operate through the price level, which is assumed to adjust so that

$$p_1(R)c_1(R) = D.$$

We do not want premature liquidation of the safe asset at date 1, so the late consumers must hold only money between dates 1 and 2. Since the nominal value of a withdrawal at date 1 is D, this implies that

$$\alpha(R)D = M.$$

Similarly, we want the early-withdrawing late consumers to be able to afford just $c_2(R)$ at date 2. To ensure this, we must have

$$\alpha(R)p_2(R)c_2(R) = M.$$

Clearly, there are many values of $\alpha(R)$, $p_1(R)$ and $p_2(R)$ that will satisfy these conditions. Furthermore, these conditions are sufficient for an equilibrium. At date 1, the bank hands out a mixture of goods and money to withdrawers. The early consumers do not want any money, so they exchange theirs with the late consumers. The late consumers do not want to hold any goods, since the return on money is greater than the return on goods:

$$\frac{p_1(R)}{p_2(R)} = \frac{c_2(R)}{c_1(R)} > 1.$$

Consequently, the late consumers end up holding only money between dates 1 and 2. At date 2, the early withdrawing late consumers supply all their money inelastically to the representative bank in exchange for goods. The representative bank gets back just enough money to repay its loan from the central bank, and has enough goods left over to give each late-withdrawing late consumer $c_2(R)$.

To see how the price level $p_1(R)$ is determined, consider the aggregate transactions at date 1. The bank gives each depositor $\frac{c_1(R)}{1+\alpha(R)}$ units of consumption and $\frac{M}{1+\alpha(R)}$ units of money. The early consumers supply a total of $\frac{M}{1+\alpha(R)}$ units of money (in exchange for goods) and the late consumers supply a total of $\alpha(R)\frac{c_1(R)}{1+\alpha(R)}$ units of goods (in exchange for money). To clear the market the price adjusts to equate the value of goods supplied to the quantity of money:

$$\frac{M}{1+\alpha(R)} = p_1(R)\alpha(R)\frac{c_1(R)}{1+\alpha(R)}$$

so

$$M = p_1(R)\alpha(R)c_1(R)$$

which is equivalent to the conditions above. The determination of $p_2(R)$ is similar.

Theorem 4 *Suppose that the central bank makes available to the representative bank an interest-free line of credit of M units of money at date* 1 *which must be repaid at date* 2. *Then there exist equilibrium price levels $p_1(R)$ and $p_2(R)$ and an equilibrium fraction of early withdrawers $\alpha(R)$, for every value of R, which will implement the incentive-efficient allocation $\{(L,X), c_1(\cdot), c_2(\cdot)\}$.*

Although the central bank policy described in Theorem 4 removes the deadweight costs of bank runs, it does not prevent the runs themselves. Injecting money into the banking system dilutes the claims of the early consumers so that they bear a share of the low returns to the risky asset. Without bank runs, incentive-efficient risk sharing would not be achieved. A policy that eliminated runs, by forcing the banking system to hold larger reserves of the safe asset, would be inefficient with respect to both risk sharing and investment.

To illustrate how the incentive-efficient allocation can be implemented in the context of the numerical example with $r = 1.05$ recall that the social optimum has $(L,X) = (1.36, 0.64)$, R = 2.34 and $\mathrm{E}[U(c_1, c_2)] = 0.32$. Suppose $D = 1.36$. For $R \geq \bar{R} = 2.34$ then $p_1(R) = p_2(R) = 1$. For $R < \bar{R} = 2.34$ the price levels at the two dates depend on the level of R. To illustrate suppose $R = 2$. In that case $c_1(2) = 1.29$ so $p_1(2) = 1.36/1.29 = 1.05$. and $c_2(2) = 1.35$ so $p_2(2) = 1.36/1.35 = 1.01$. Similarly for other values of *R*. Note that it is optimal at these prices for the early withdrawers to hold money from date 1 to date 2 since the price of goods is falling. In other words, they won't use the storage technology available to them because they can do better holding money. The fraction of late consumers who withdraw from the bank and hold money will be determined by M. Suppose $M = 1$. Then $\alpha(R) = 1/1.36 = 0.74$.

IV. Asset Trading and the Efficiency of Runs

As has been pointed out above, the total illiquidity of the long-term, risky asset plays an important role in equilibrating bank runs, so that runs are typically partial, that is, involve only a fraction of the late consumers. Introducing an asset market, thus allowing the bank to liquidate its holding of the risky asset by selling it on the market, has a number of implications. In the first place, it allows the bank to use all its assets to meet the demands of the early withdrawers, assuming this is required by the terms of the deposit contract. In the second place, the possibility of liquidating the risky asset is likely to make the bank run worse. This happens in two ways. First, if the market for the risky asset is at all illiquid, the sale of the representative bank's holding of the risky asset will drive down the price, thus making it harder to meet the depositors' demands. Secondly, because a bank run exhausts the bank's assets at date 1, a late consumer who waits until date 2 to withdraw will be left with nothing. So whenever there is a bank run it will involve all the late consumers and not just some of them.

The all-or-nothing character of bank runs is, of course, familiar from the work of Diamond and Dybvig (1983). The difference here is that bank runs are assumed to occur only when there is no other equilibrium outcome possible. Furthermore, the deadweight cost of a bank run in this case is endogenous. In addition to the explicit cost of liquidation ($r > 1$), there is a cost resulting from suboptimal risk sharing. To make this clear, in this section we assume that $r = 1$. When the representative bank is forced to liquidate the risky asset, it sells the asset at a low price. This is a transfer of value to the purchasers of the risky asset, not an economic cost. The deadweight loss arises because the transfer occurs in bad states when the consumers' consumption is already low. In other words, the market is providing negative insurance.

Once again, intervention by the central bank will be helpful, but the optimal policy will consist of eliminating the deadweight costs of runs that arise from premature liquidation, rather than eliminating the runs themselves.

A. The Asset Market

To make these ideas precise, we assume that there are two classes of agents, the risk averse consumers, who use the banking system to make investments for them, and a group of risk neutral *speculators,* who make direct investments in the safe and risky assets. Speculators consume only in the last period and their objective is to maximize the expected value of their portfolio at date 2. The speculators are all identical, so they can be replaced by a representative individual, who has an initial wealth W_s and chooses a portfolio $(L_s, X_s) \geq 0$ subject to the budget constraint

$L_s + X_s = W_s$. The assumption that holdings of the two assets must be non-negative is important here. Risk neutrality is often interpreted as meaning that an individual can have unboundedly negative consumption and hence supply unboundedly large amounts of the liquid asset. Such an interpretation would make no sense here, because we want to emphasize the consequences of restricted liquidity in the market.

Since the risky asset has a higher expected return than the safe asset, the safe asset will be held only if the speculators can make a profit by buying the risky asset at a low price at date 1. If bank runs occur in equilibrium, the safe asset must be held by speculators. If speculators do not have a positive holding of the safe asset at date 1, then when the banks try to sell the risky asset the price will fall to zero in some states, which means that any speculator who had held the safe asset would make an infinite profit. (Note the importance for this argument of the assumption that speculators cannot short the safe asset). Thus, in an equilibrium where runs occur with positive probability, $L_s > 0$.

On the other hand, if W_s is large enough (as we assume in the sequel) speculators must also hold the risky asset. If not, $L_s = W_s$. and if the price of the risky asset is less than its "fair" value R at date 1, this amount of the safe asset will be supplied in exchange for the amount of the risky asset offered by the banks. Since $X \leq E$, the price must be at least W_s/E. So the speculators only make a profit if $R > W_s/E$. However, as we shall see, the banks only sell the risky asset when the return R is sufficiently small, so by choosing W_s large enough we can ensure that the speculators profit only if $L_s < W_s$. To sum up, there is no loss of generality in assuming that $L_s > 0$ and $X_s > 0$ in any equilibrium in which bank runs occur with positive probability.

The necessary and sufficient condition for holding both assets to be an optimum for the speculator is that

$$\mathrm{E}[\max\{1, \frac{R}{P(R)}\}] = \mathrm{E}[R]$$

where $P(R)$ is the price of the risky asset at date 1. In other words, the expected return from holding the safe asset and buying the risky asset at date 1 when the price of the risky asset falls below R is equal to the expected return from a buy-and-hold strategy, that is buying the risky asset at date 0 and holding it until date 2. Note that $P(R) \leq R$ for all values of R, because $P(R) > R$ implies that no one is willing to hold the risky asset and this cannot be an equilibrium. Therefore, we do not have to consider the possibility of switching from the risky to the safe asset at date 1 and the condition above reduces to

$$E[\frac{1}{P(R)}] = E[R] \ (*)$$

B. The Bank's Decision

The bank chooses a portfolio (L, X) and a promised payout $\bar{c}$ at date 1 subject to the usual feasibility and incentive constraints. The standard deposit contract requires the bank to pay an early withdrawer $\bar{c}$ at date 1, if this is possible, and to liquidate all its assets otherwise. If there is no run, the early consumers receive $c_1(R) = \bar{c}$ and late consumers receive whatever is left, i.e., $c_2(R) = L - \bar{c} + RX$. We do not want to allow a run unless it is unavoidable: when there are multiple equilibria corresponding to a given value of R, we assume that the equilibrium without runs is chosen. There are two possible cases to consider. A run will occur if and only if it is impossible to pay the early consumers $\bar{c}$ and pay the late consumers an amount at least as great as $\bar{c}$. If there is a run, the bank must liquidate all its assets at date 1 and all late consumers will join the run. In that case, $c_1(R) = c_2(R) = \frac{1}{2}(L + P(R)X)$.

Let R^* be implicitly defined by the condition

$$\bar{c} = \frac{1}{2}(L + R^* X).$$

Then a run occurs if and only if $R < R^*$. To see this, suppose that $R < R^*$. If there is no run, then $c_1(R) + c_2(R) \geq 2\bar{c} > L + RX$, contradicting the feasibility conditions. (Since $P(R) \geq R$ selling assets will not help either). On the other hand, if $R \geq R^*$ then it is clearly possible to choose $c_1(R) = \bar{c}$ and $c_2(R) = L - \bar{c} + RX \geq \bar{c}$.

The bank's decision problem can be written as follows:

$$
(P5) \qquad
\begin{array}{lll}
\max & & \mathrm{E}[u(c_1(R) + u(c_2(R))] \\
\text{s.t.} & \text{(i)} & L + X \leq E; \\
 & \text{(ii)} & c_1(R) = \begin{cases} \bar{c} & \text{if } R \geq R^* \\ \frac{1}{2}(L + P(R)X) & \text{if } R < R^* \end{cases}; \\
 & \text{(iii)} & c_2(R) = \begin{cases} L - \bar{c} + RX & \text{if } R \geq R^* \\ \frac{1}{2}(L + P(R)X) & \text{if } R < R^* \end{cases};
\end{array}
$$

where $R^* = (2\bar{c} - L)/X$.

C. Equilibrium

An equilibrium for the model with an asset market consists of a portfolio (X_s, L_s) and a price function $P(R)$ that satisfies the no-arbitrage condition $(*)$, and a deposit contract $((L, X), \bar{c})$ that solves the decision problem *(P5)* given the values of (X_s, L_s) and $P(R)$.

In the asset market, our earlier discussion shows that there are two cases to be considered: either $R \geq R^*$ and there is no run or $R < R^*$ and there is a run. If there is no run and hence no sale of assets in the market, the safe asset must have the same one-period return as the risky asset, so $P(R) = R$. On the other hand, if there is a sale of assets, the representative bank supplies X inelastically. If $L_s > RX$, then the equilibrium price must be $P(R) = R$. If the price were lower, everyone would want to hold the risky asset and there would be an excess supply of the safe asset. If the price were higher, no one would want to hold the risky asset and there would be an excess supply of the risky asset. On the other hand, if $L_s < RX$ then the price of the risky asset must be $P(R) = L_s/X$. At this price, the speculators supply the safe asset inelastically in exchange for the risky asset and the market clears because $L_s = P(R)X$. At any other price, this market-clearing condition will be violated. (If $P(R) = R$, speculators may supply less than L_s, but this too violates market clearing). Let R^0 be implicitly defined by the condition

$$L_s = R^0 X.$$

Then

$$P(R) = \begin{cases} R & \text{for } R < R^0 \text{ and } R \geq R^* \\ L_s/X & \text{for } R^0 < R < R^*. \end{cases}$$

In other words, the price collapses only if the return is low enough to provoke a run but not so low that the market is liquid enough to absorb the asset at its "fair" value. Figure 5 illustrates the equilibrium allocation for bank depositors.

In the numerical example it will be assumed that the wealth of the speculators $W_s = 1$ and that the other parameters are the same as in the standard case with $r = 1$. The optimal contract for depositors has $(L, X) = (1.06, 0.94)$, $R^0 = 0.25, R^* = 1.13$, with $P(R) = 0.25$ for $R^0 < R < R^*$ and $\text{E}[U(c_1, c_2)] = 0.09$. For the speculators $(L_s, X_s) = (0.24, 0.76)$ and their expected utility is $\text{E}[U_s] = 1.5$. Note that the depositors are significantly worse off in this equilibrium compared to the incentive-efficient allocation (*P1*) where $\text{E}[U(c_1, c_2)] = 0.25$ and are only slightly better off than in the case where the bank's portfolio is such that no runs occur (as in Figure 2) in which case $\text{E}[U(c_1, c_2)] = 0.08$.

D. Optimal Policy

When we come to analyze the possibilities for welfare-improving monetary intervention, it is not immediately clear how to proceed. The existence of risk-neutral speculators obviously gives rise to the potential for risk sharing that is not being provided by the market. For example, if the speculators assumed more of the risk associated with the risky asset, the risk-averse depositors would clearly be better off and yet there is no way that the simple asset market at date 1 would be able to accomplish this allocation of risk. This does not seem a very interesting benchmark by which to judge the market allocation. It goes beyond what we normally think a central bank can achieve and it assumes an ability on the part of the central bank to enforce contingent contracts that we have assumed are too costly for the market. Since it is not clear why the central bank would have this advantage over the market, this seems an excessively strong standard by which to judge the market.

On the other hand, if we assume that the central bank can only trade on the asset market like the representative bank, there is another problem. The speculators will only hold the safe asset if there is a positive probability that they will be able to make profits by buying the risky asset at less than its "fair" value at date 1. Ex post, the central bank will be able to control the price of the risky asset by choosing to supply the revenue-maximizing amount in each state. It is easy to see that the revenue-maximizing amount of the asset supplied will always correspond to a price equal to the "fair" value of the asset, i.e., $P(R) = R$. Consequently, unless the central bank can commit to a pricing policy in advance and thus eliminate the time-inconsistency problem, the speculators will have no incentive to hold the safe asset and the asset market will not be useful for obtaining additional liquidity.

Rather than pursuing these issues here, we choose as a benchmark the allocation that solves *(P1)*. This allocation can be implemented without relying on the asset market at all. It may not be the best the central bank can do, whatever one chooses to define as the "best", but it provides a lower bound for the second best and for some parameter values we can show that it is significantly better than the equilibrium allocation. The essential idea behind the policy that implements the solution to *(P1) is* similar to the monetary intervention described in Section 3, but here the central bank is interpreted as supporting the risky asset's price, rather than making an unsecured loan to the bank. Specifically, the central bank enters into a repurchase agreement (or a collateralized loan) with the representative bank, whereby the bank sells some of its assets to the central bank at date 1 in exchange for money and buys them back for the same price at date 2. By providing liquidity in this way, the central bank ensures that the representative bank does not suffer a loss by liquidating its holdings of the risky asset prematurely.

As before, we assume that the standard deposit contract promises depositors a fixed amount of money D in the middle period and pays out the remaining value of the assets in the last period. The price level at date t in state R is denoted by $p_t(R)$ and the *nominal* price of the risky asset at date 1 in state R is denoted by $P(R)$. We want the risky asset to sell for its "fair" value, so we assume that $P(R) = p_1(R)R$. At this price, the safe and risky assets are perfect substitutes. Let (X, L) be the portfolio corresponding to the solution of *(P1)* and let $(c_1(R), c_2(R))$ be the corresponding consumption allocations. For large values of R, we may have $c_1(R) = L < c_2(R) = RX$; for smaller values we may have $c_1(R) = c_2(R) = \frac{1}{2}(L + RX)$. To implement this allocation requires introducing contingencies through price variation: $p_1(R)c_1(R) = D < p_2(R)c_2(R)$ for $R > \bar{R}$ and $p_1(R)c_1(R) = D = p_2(R)c_2(R)$ for $R < \bar{R}$. These equations uniquely determine the value of $p_1(R)$ and the value of $p_2(R)$ when $R < \bar{R}$. It remains only determine the value of sales of assets and the size of the bank run.

In the event of a bank run, only the late consumers who withdraw early will end up holding cash, since the early consumers want to consume their entire liquidated wealth immediately. If $\alpha(R)$ is the fraction of late consumers who withdraw early, then the amount of cash injected into the system must be $\alpha(R)D$. For simplicity, we assume that the amount of cash injected is a constant M and this determines the "size" of the run $\alpha(R)$. Since the safe asset and the risky asset are perfect substitutes at this point, it does not matter which assets the representative bank sells as long as the nominal value equals M. The representative bank enters into a repurchase agreement under which it sells assets at date 1 for an amount of cash equal to M and repurchases them at date 2 for the same cash value.

At the prescribed prices, speculators will not want to hold any of the safe assets, so $L_s = 0$ and $X_s = W_s$.

It is easy to check that all the equilibrium conditions are satisfied: depositors and speculators are behaving optimally at the given the prices and the feasibility conditions are satisfied.

Theorem 5 *The central bank can implement the solution to problem (P1) by entering into a repurchase agreement with the representative bank at date* 1. *Given the allocation* $\{(L, X), c_1(R), c_2(R)\}$, *corresponding to the solution of* $(P1)$, *the equilibrium values of prices are given by the conditions* $p_1(R)c_1(R) = D < p_2(R)c_2(R)$ *for* $R > \bar{R}$ *and* $p_1(R)c_1(R) = D = p_2(R)c_2(R)$ *for* $R < \bar{R}$. *There is a fixed amount of money M injected into the economy in the event of a run and the fraction of late withdrawers who "run" satisfies* $\alpha(R)D = M$. *The price of the risky asset at date* 1 *satisfies* $p_1(R)R = P(R)$ *and the optimal portfolio of the speculators is* $(L_s, X_s) = (0, W_s)$.

While Theorem 5 shows the central bank intervention can achieve the planner's solution to *(P1)*, it does not show that this is better than the market equilibrium, since the market equilibrium allows for possibilities, such as liquidating the risky asset at date 1, which are not available in (P1). However, it is easy to show that the solution to *(P1)* is Pareto-preferred to the equilibrium of the model with asset markets. To see this, let (X_s, L_s) be the speculators' equilibrium portfolio, $P(R)$ the equilibrium asset-price function, and $\{(L, X), c_1(R), c_2(R)\}$ the equilibrium deposit contract. The consumption functions solve

$$\begin{aligned} \max \quad & \mathrm{E}[u(c_1(R)) + u(c_2(R))] \\ \text{s.t.} \quad & c_1(R) \le L \\ & c_1(R) + c_2(R) \le L + RX \end{aligned}$$

if $R \ge R^*$ and

$$\begin{aligned} \max \quad & \mathrm{E}[u(c_1(R)) + u(c_2(R))] \\ \text{s.t.} \quad & c_1(R) + c_2(R) \le L + P(R)X \end{aligned}$$

if $R < R^*$. Note that $c_1(R) \le \bar{c} \le L$ for all values of R, so there is no loss of generality in combining these two problems and treating $(c_1(R), c_2(R))$ as the solution of

$$\begin{aligned} \max \quad & \mathrm{E}[u(c_1(R)) + u(c_2(R))] \\ \text{s.t.} \quad & c_1(R) \le L \\ & c_1(R) + c_2(R) \le L + \min\{P(R), R\}X \end{aligned}$$

for all values of R.

Now suppose that the functions $c_1^*(R)$ and $c_2^*(R)$ solve the problem

$$\begin{aligned} \max \quad & \mathrm{E}[u(c_1(R)) + u(c_2(R))] \\ \text{s.t.} \quad & c_1(R) \le L \\ & c_1(R) + c_2(R) \le L + RX \end{aligned}$$

for all values of R. Then since $P(R) \le R$ we must have $c_i(R) \le c_i^*(R)$, for all R and $i = 1, 2$. The consumption functions $c_1^*(R)$ and $c_2^*(R)$ are feasible for $(P1)$ if the portfolio (L, X) is chosen, so it follows that the solution to *(P1)* must be at least as good as the equilibrium outcome and strictly preferred by the depositors if the equilibrium involves selling the risky asset at a price $P(R) < R$ with positive probability.

The speculators get the same expected utility in either case, so we have the following result.

Corollary 5.1 *The solution to (P1), implemented by the policy described in Theorem 5, is Pareto-preferred to the laisser-faire equilibrium outcome of the model with asset markets.*

Theorem 5 and its corollary can be illustrated with the standard numerical example. To illustrate how the incentive-efficient allocation *(P1)* can be implemented in the context of the numerical example with $r = 1$ recall that the optimum has $(L, X) = (1.19, 0.81)$, $\bar{R} = 1.47$ and $\mathrm{E}[U(c_1, c_2)] = 0.25$. Suppose $D = 1.19$. For $R \geq \bar{R} = 1.47$ then $p_1(R) = p_2(R) = 1$. For $R < \bar{R} = 1.47$ the price levels at the two dates depend on the level of R. To illustrate suppose $R = 1$. In that case $c_1(1) = c_2(1) = 1$ so $p_1(1) = p_2(1) = 1.19$. Similarly for other values of R. The lower the value of R the higher $p_t(R)$ so that consumption is lowered by raising the price level. Also $P(R) = 1.19$. The fraction of late consumers who withdraw from the bank and hold money will be determined by M. Suppose $M = 1$ then $\alpha(R) = 1/1.19 = 0.84$. For the speculators $(L_s, X_s) = (0, 1)$ and their expected utility is $\mathrm{E}[U_s] = 1.5$. The equilibrium with central bank intervention is clearly Pareto-preferred to the market equilibrium without intervention as indicated by the corollary.

E. Equilibrium without Runs

Because of the inefficiency of an equilibrium with bank runs when there is no monetary intervention, it may be better to have no runs, even if this means holding an inefficiently high level of reserves. To rule out runs entirely, without recourse to the asset market, the bank must hold enough reserves in the form of the safe asset to guarantee late consumers the same consumption as early consumers for every value of R. If the minimum value of R is zero, this means that L must be at least $2\bar{c}$. In that case, the consumption of the early consumers is always $\bar{c}$ and consumption of the late consumers is $RX + L - \bar{c} \geq \bar{c}$. The bank chooses the portfolio to maximize $\mathrm{E}[u(\bar{c}) + u(L - \bar{c} + RX)]$ subject to the constraint $2\bar{c} \leq L$ and, of course, $L + X \leq E$. For example, in the standard example $\mathrm{E}[U(c_1, c_2)] = 0.09$ in the market equilibrium but 0.08 when banks choose a portfolio which eliminates runs. Changing the example slightly by replacing the assumption that the probability density function of R is uniform on [0,3] with the assumption it is uniform on [0, 2.9] but keeping everything else the same leads to a situation where banks will voluntarily choose an allocation $(L, X) = (1.66, 0.34)$. The optimal deposit contract is as in Figure 2 and runs do not occur.

However, if the banks choose to hold a portfolio that is inconsistent with runs, there is no need to have a policy that imposes this solution, and if they do not choose such a portfolio, imposing one by regulation will make depositors strictly worse off. Furthermore, even if banks choose to hold a very safe portfolio in equilibrium because of the costs of runs, it does not follow that the optimal policy which implements the incentive-efficient will not involve runs. On the contrary, we have seen that for a plausible parametric specification, runs will be an integral part of the optimal policy, regardless of the presence or absence of runs in a laisser-faire equilibrium. In the example where the probability density function is uniform on [0, 2.9] the incentive-efficient allocation is similar to that in Figure 1 and involves runs.

V. Concluding remarks

Empirical evidence provided by Gorton (1988) suggests that banking panics in the U.S. during the National Banking Era were not sunspot phenomena but rather were the result of the business cycle. When depositors observe leading economic indicators and perceive that a bank's receipts are going to be low there is a run. This paper has developed a simple model of this type of run and used it to identify the benefits and costs of runs. It has been shown that financial crises can be optimal if the return on the safe asset is the same inside and outside of the banking system. The reason is that the optimal allocation of resources often involves investing a significant amount in risky assets and imposing some risk on people who withdraw early. Allowing bank runs can be an efficient way of doing this. In this case central bank policies and actions of other government agencies, which eliminate runs, can lower the welfare of depositors. On the other hand, if the return on the safe asset is higher inside the banking system than outside so bank runs are costly runs alone cannot achieve the optimal allocation of resources. However, a monetary intervention by central bank can allow the first best to be achieved. Finally, if the risky asset can be sold in an asset market, bank runs may be costly even when the return on the safe asset is the same inside and outside the banking system. The reason is that banks are forced to liquidate their asset when prospects are bad. This simultaneous liquidation drives the price down and allows speculators in the asset market to profit. There is, in effect, negative insurance. Central bank intervention which prevents the collapse in prices in the asset market can allow a Pareto improvement.

The assumption of a representative bank in our model means that the prospect of poor returns on the risky asset causes an economy wide effect. This precludes the consideration of a number of interesting features of actual panics. The first is the *fragility* of the banking system. It is often argued that bank failures are likely to spread by *contagion*. Our model would need to be extended to include heterogeneous banks to articulate a theory of banking panics (contagions). One of the most important effects of bank runs is the (possibly permanent) closure of the affected banks, increasing the costs of intermediation for the entire economy. In extreme cases, financial disruption in the banking system may have a severe effect on aggregate economic activity (Bernanke (1983)). This effect can be captured in a dynamic model with bank capital. Suppose there are two classes of agents, those with large wealth, who become bankers, and those with small wealth, who become depositors. Bank failures result in the transfer of wealth from bankers to depositors. In subsequent periods, there is less bank capital available, the cost of intermediation will be higher, and the economy will be poorer. This will be true even if the assets themselves have not been destroyed and there has been no change in the banking technology. The agency costs of providing capital to the banking system makes the efficiency of banking services dependent on the distribution of wealth (Bernanke and Gertler (1989)). When these dynamic effects of bank runs are taken into account, there may be an additional reason for intervention. In effect, it is protecting the capital of the banking sector. In the short-run this may appear to be at the expense of depositors, but in the long run, even depositors may be better off. Of course, the depositors could in principle make transfers to the banking sector, but the free-rider problem makes this impractical.

Appendix

Proof of Theorem 1

If we ignore the incentive-compatibility constraint, the optimal risk-sharing problem becomes:

$$(P1)' \quad \begin{array}{lll} \max & & E[u(c_1(R)) + u(c_2(R))] \\ \text{s.t.} & \text{(i)} & L + X \leq E; \\ & \text{(ii)} & c_1(R) \leq L; \\ & \text{(iii)} & c_1(R) + c_2(R) \leq L + RX. \end{array}$$

A necessary condition for a solution to $(P1)'$ is that, for each value of R, the consumption levels $c_1(R)$ and $c_2(R)$ solve the problem

$$\begin{array}{lll} \max & & u(c_1(R)) + u(c_2(R)) \\ \text{s.t.} & \text{(ii)} & c_1(R) \leq L; \\ & \text{(iii)} & c_1(R) + c_2(R) \leq L + RX. \end{array}$$

The necessary Kuhn-Tucker conditions imply

$$u'(c_1(R)) \geq u'(c_2(R))$$

with strict equality if $c_1(R) < L$. In any case, this implies that $c_1(R) \leq c_2(R)$, with strict equality if $c_1(R) < L$, so the incentive constraints (iv) will be satisfied automatically. Thus, a solution to $(P1)'$ is also a solution to the original problem $(P1)$.

Since we know that $c_1(R) = c_2(R)$ whenever $c_1(R) < L$, there are two regimes to be considered. Either $c_1(R) = L$ and (hence) $c_2(R) = RX$ or $c_1(R) = c_2(R) = \frac{1}{2}(RX + L)$. The first case can occur if and only if $L \leq RX$, so the optimal risk-sharing allocation must satisfy

$$c_1(R) = c_2(R) = \frac{1}{2}(RX + L) \text{ if } L \geq RX$$

and

$$c_1(R) = L, c_2(R) = RX \text{ if } L \leq RX.$$

This allows us to write the optimal risk-sharing problem more compactly as follows:

$$\max \int_0^{\bar{R}} 2u\left(\frac{RX+L}{2}\right) f(R)dR + \int_{\bar{R}}^{\infty} (u(L) + u(RX)) f(R)dR$$

$$\text{s.t. } L + X \leq E,$$

where $\bar{R} \equiv L/X$ is the value of the return on the risky asset at which the liquidity constraint begins to bind. Note that so far we have not established that the critical value of R belongs to the support of R.

It remains to characterize the optimal portfolio. We first rule out two extreme cases. Suppose that $X = 0$. Then it is clear that $c_1(R) = c_2(R) = E/2$ and $\bar{R} = \infty$. This will be optimal only if $L = E$ maximizes

$$u(L/2) + \mathrm{E}[u(R(E-L) + L/2)]$$

and the first-order condition for this is

$$u'(E/2)/2 + u'(E/2)(\frac{1}{2} - \mathrm{E}[R]) \geq 0,$$

which implies $\mathrm{E}[R] \leq 1$, contradicting one of our maintained assumptions.

Next suppose that $L = 0$. Then $c_1(R) = 0 \leq c_2(R) = RE$. For this to be an optimal choice, it must be the case that X = E maximizes

$$u(E-X) + \mathrm{E}[u(RX)]$$

and the necessary first-order condition for this is

$$u'(0) \leq \mathrm{E}[u'(RE)R]$$

which contradicts another of our maintained assumptions. Thus any optimal portfolio must satisfy $L > 0$ and $X > 0$.

Returning to the compact form of the risk-sharing problem above, we see that a necessary condition for an interior solution is:

$$\int u'(c_1(R))f(R)dR = \lambda$$

and

$$\int u'(c_2(R))Rf(R)dR = \lambda$$

where λ is the Lagrange multiplier of the constraint $L + X = E$. Under the strict concavity of $u(\cdot)$, these first-order conditions uniquely determine the optimal values of L and X, which in turn determine $\bar{R}$, $c_1(R)$, and $c_2(R)$ through the relationships described above. ■

References

Bernanke, B. (1983). "Non-monetary Effects of the Financial Crisis in the Propagation of the Great Depression," *American Economic Review* 73, 257-263.

Bernanke, B. and M. Gertler (1989). "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review* 79, 14-31.

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

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

Gorton, G. (1988). "Banking Panics and Business Cycles," *Oxford Economic Papers* 40, 751-781.

Kindleberger, C. P. (1978). *Manias, Panics, and Crashes: A History of Financial Crises,* New York: Basic Books.

Mitchell, W. C. (1941). *Business Cycles and Their Causes,* Berkeley: University of California Press.

Postlewaite, A. and X., Vives (1987). "Bank Runs as an Equilibrium Phenomenon," *Journal of Political Economy* 95, 485-491.

Waldo, D. (1985). "Bank Runs, the Deposit-Currency Ratio and the Interest Rate," *Journal of Monetary Economics* 15, 269-278.

Table 1

National Banking Era Panics

NBER Cycle Peak–Trough	Panic Date	%Δ(Currency/ Deposit)*	%Δ Pig Iron†
Oct. 1873–Mar. 1879	Sep. 1873	14.53	−51.0
Mar. 1882–May 1885	Jun. 1884	8.80	−14.0
Mar. 1887–Apr. 1888	No Panic	3.00	−9.0
Jul. 1890–May 1891	Nov. 1890	9.00	−34.0
Jan. 1893–Jun. 1894	May 1893	16.00	−29.0
Dec. 1895–Jun. 1897	Oct. 1896	14.30	−4.0
Jun. 1899–Dec. 1900	No Panic	2.78	−6.7
Sep. 1902–Aug. 1904	No Panic	−4.13	−8.7
May 1907–Jun. 1908	Oct. 1907	11.45	−46.5
Jan. 1910–Jan. 1912	No Panic	−2.64	−21.7
Jan. 1913–Dec. 1914	Aug. 1914	10.39	−47.1

*Percentage change of ratio at panic date to previous year's average.

†Measured from peak to trough.

(Adapted from Table 1, Gorton (1988), p. 233.)

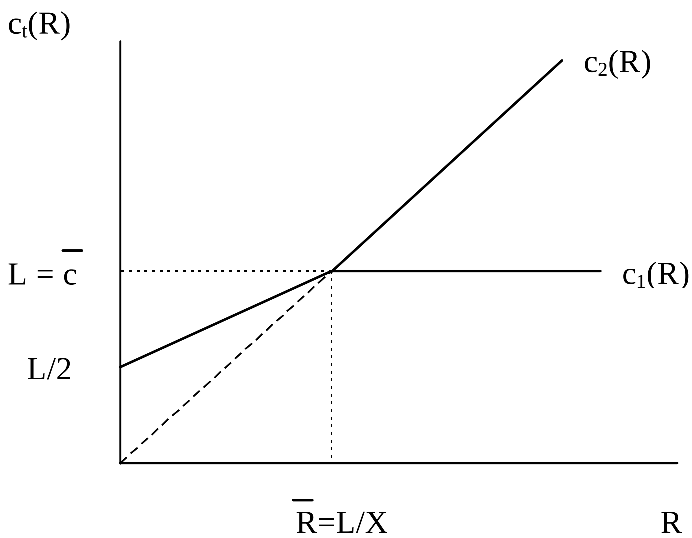

Figure 1

The optimal risk sharing allocation and the optimal deposit contract with runs

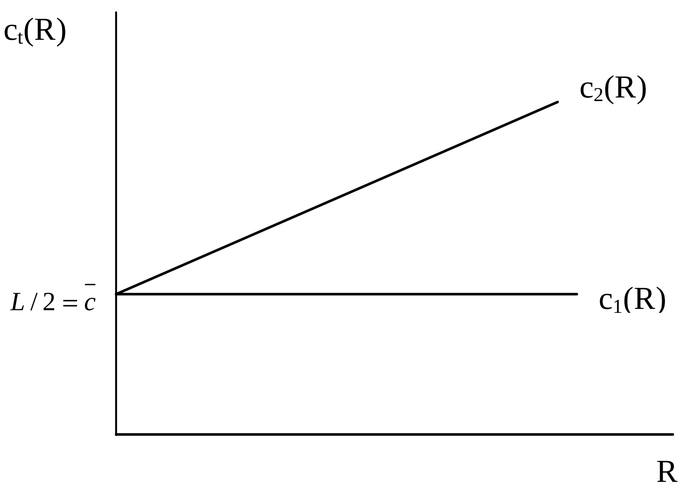

Figure 2

The optimal deposit contract without runs

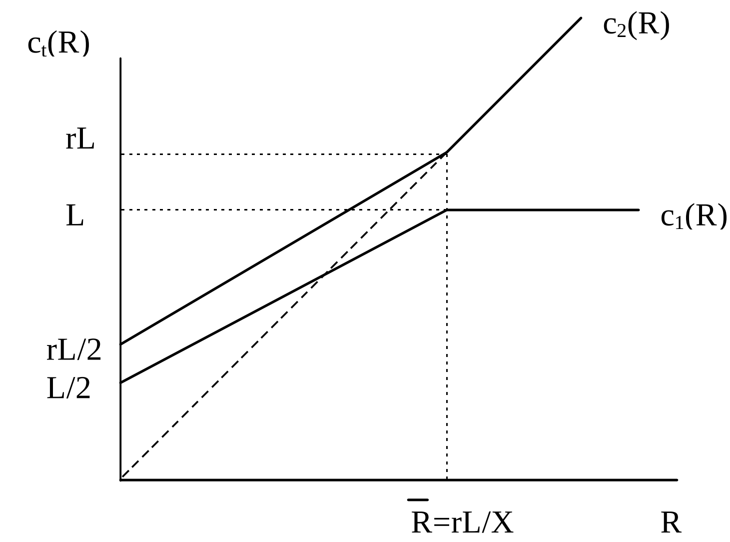

Figure 3

The optimal risk sharing allocation with costly liquidation

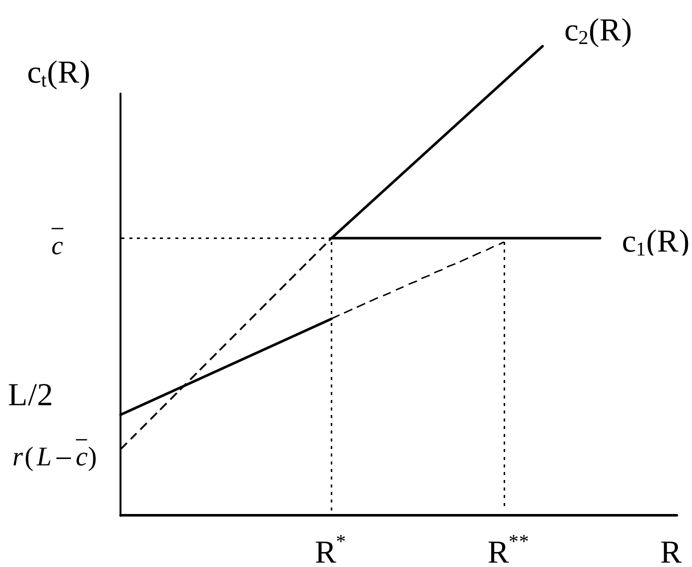

Figure 4

The optimal deposit contract with costly liquidation when $\bar{c} < L$

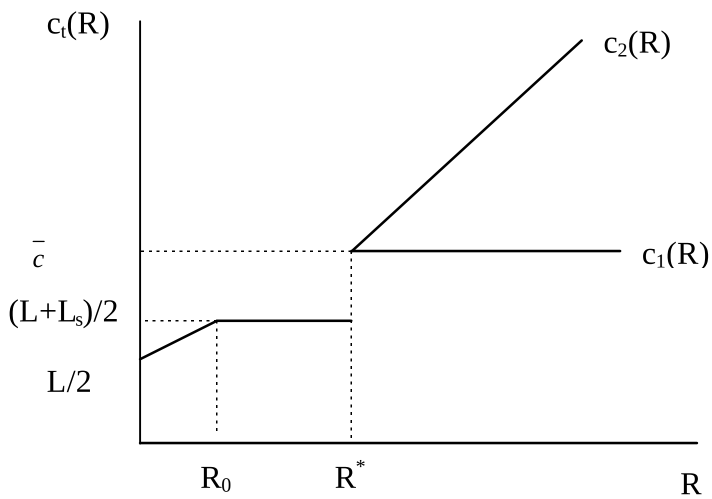

Figure 5

The optimal deposit contract when there is a market for the risky asset