Financial Fragility in Monetary Economies*

David Andolfatto
Federal Reserve Bank of St. Louis and *Simon Fraser University*

Aleksander Berentsen
University of Basel

Fernando M. Martin
Federal Reserve Bank of St. Louis

February 15, 2016

Abstract

This paper integrates the Diamond and Dybvig (1983) theory of financial fragility with the Lagos and Wright (2005) model of monetary exchange. Non-bank monetary economies with well-functioning secondary markets for capital can allocate risk reasonably well, but are never efficient. When secondary markets are subject to "market freeze" events, risk-sharing deteriorates accordingly. A fractional-reserve bank can dominate a monetary economy because: (i) it provides superior risk-sharing even when market freeze events are absent; and (ii) it bypasses the need for a secondary capital market to begin with. Indeed, a fractional reserve bank can implement the optimal allocation when monetary policy follows the Friedman rule. However, the desirability of fractional reserve banking is diminished if the structure is subject to "bank run" events. In the event of a run, an open secondary market allows banks to liquidate capital at a price that permits honoring all deposit obligations. If bank runs are expected to occur with a sufficiently high probability, then a narrow banking structure may be preferred. Narrow banks are more stable, but offer less risk-sharing. We find that the choice of bank structure can depend on monetary policy. High inflation economies penalize narrow banking systems relatively more than fractional reserve systems. In this way, high inflation promotes financial instability. We also show how special interests are not generally aligned over the choice of bank regime.

*We thank Jean-Charles Rochet and Steve Williamson for several very useful comments and criticisms.

1 Introduction

This paper integrates the Diamond and Dybvig (1983) theory of financial fragility with the Lagos and Wright (2005) model of monetary exchange. The primary purpose of this exercise is to discover whether monetary policy affects the incentives for agents to form fragile banking systems, whether inflation-financed lender-of-last-resort facilities can promote economic efficiency, and whether monetary policy geared toward lender-of-last-resort interventions are likely to favor some special interests over others.

Our model economy is populated by investors and workers. Investors are endowed with a sequence of illiquid capital projects that yield a high expected return if held to maturity, but which earn a low risk-free return if liquidated early. Investors are subject to random, but insurable, expenditure needs over time. In any given period, a known fraction of investors will need funds prior to the time a capital project matures. The timing of investor funding needs is private information.

Workers in the model provide the services wanted by investors with early expenditure needs. Workers are unwilling to accept private investor debt as payment for labor services.[1] As a consequence, wages must be paid in cash, which we assume here to take the form of central bank liabilities. As in Diamond and Dybvig (1983), investors are motivated to enter a risk-sharing arrangement that permits them to enjoy the fruits of high return capital with the flexibility of a liquid payment instrument. The optimal risk-sharing arrangement entails the creation of a fractional reserve bank–that is, an agency that acquires assets consisting of cash reserves and capital investments funded through deposit liabilities made redeemable on demand for cash.

We assume that banks adopt the standard bank deposit contract considered by Diamond and Dybvig (1983). That is, implicitly, banks cannot condition payments on rumors of an impending run.[2] When this is so, the standard deposit contract is run-prone, that is, rumors of an impending run can become a self-fulfilling prophecy. The existence of multiple equilibria implies that an economy is "fragile" in the sense that it is prone to coordination failure. In the present context, even investors that would not normally make cash withdrawals are induced to do so if they believe that others will act likewise. Because a fractional reserve bank cannot honor all of its short-term obligations when early redemption requests are made *en masse*, it is forced to liquidate its assets at a significant loss.

Investors in our model are faced with a trade-off. On the one hand, fractional reserve banking confers benefits: a liquid, high-return payment instrument. On the other hand, the liquidity mismatch between bank assets and liabilities opens the door to inefficient liquidation events. But investors also have an option to create a run-proof financial system. They can do so, for example, by insisting that banks back all demandable liabilities fully

[1] Assume, for example, that investors cannot commit to the promises they make to workers and that investor-worker relationships are difficult to form.

[2] As is well-known, if banks could credibly threaten to suspend redemptions once cash reserves are depleted, then the standard deposit contract can prevent sunspot-driven bank runs in the Diamond and Dybvig (1983) model. Andolfatto, Nosal and Sultanum (2014) demonstrate that there exist indirect mechanisms that can eliminate bank run equilibria in much more general environments characterized by sequential service and aggregate uncertainty.

with cash reserves. The stability of this "narrow banking" regime, however, comes at the cost of lower funding for capital investments. The relative net benefits of these two regimes depends on, among other things, the propensity among citizens to succumb to infectious rumors leading to coordination failure. We treat this propensity an exogenous parameter describing an inherent social trait that we label *social* or *inherent fragility*. Formally, we model it as the probability of a "sunspot" that triggers the rumor of a run.

Our model predicts, *ceteris paribus*, that investors living in fragile societies prefer to construct relatively stable financial systems. If fractional reserve banking systems were to exist in such societies, bank runs and financial crises would occur at high frequency–something we do not observe. Conversely, our model predicts that financial systems with a manufactured *financial fragility* (fractional reserve banking regime) are likely to emerge only in economies that are not socially fragile. It follows as a corollary that bank runs and financial crises are likely to be low frequency events, which is also consistent with the evidence.

A distinguishing characteristic of our model is that deposit liabilities are redeemable on demand for cash instead of goods. For many questions of interest, such a distinction may not be terribly important. The key insight of Diamond and Dybvig (1983), for example, is independent of what exactly constitutes the object of redemption. But monetary policy is usually conducted through a central bank and a central bank is, after all, just a special type of bank. As such, one would expect central banks to play a special role in shaping the structure and behavior of a financial system.[3]

We begin first by asking how high and low inflation rate regimes are likely to affect the banking sector. Our model predicts that inflation has only a modest impact on the size of the banking sector's balance sheet, but a more significant impact on the composition of its assets. Because high inflation regimes penalize zero-interest cash reserves, high inflation induces banks to economize on cash reserves and expand their loan portfolios. Because narrow banks have larger cash holdings than fractional reserve banks, narrow banking systems suffer relatively more under high inflation regimes. The implication of this is that narrow banks that were previously at the margin are induced to convert to fractional reserve banks in an attempt to economize on cash reserves. In short, high inflation induces manufactured financial fragility.

Next, we investigate the costs and benefits of a central bank funded lender-of-last-resort (LOLR) facility. A central bank is in a unique position, of course, to help private banks honor their short-term obligations. We model the collateral that banks offer for their emergency loans as risky. When the collateral turns out to be worthless, some investors become *de facto* recipients of helicopter money. On the one hand, the LOLR facility has the benefit of reducing (possibly eliminating) the incidence of bank runs. But on the other hand, the extent to which transfers to the banking sector are financed through money creation, the LOLR creates a distortionary expected inflation. Numerical examples suggest that these costs are small relative to their benefits, at least, as far as investors are concerned. In short, the model suggests that investors will lobby hard for fractional reserve banking systems supported by an inflation-financed LOLR facility.

[3]This is not meant to downplay the importance of fiscal and regulatory policies, of course.

It is interesting to report, however, that the workers in our model do not necessarily have their interests aligned with investors. As it turns out, the economic welfare that workers enjoy is proportional to the aggregate demand for real cash balances. The switch from narrow to fractional reserve banking has the effect of, among other things, reducing the demand for real cash balances. Moreover, monetary transfers to investors through the banking system also have the effect of redistributing purchasing power away from workers toward investors. Our work in this area remains preliminary.

2 Related literature

There is, of course, a long history of economic analysis concerned with the causes and consequences banking panics.[4] As Gorton (1988) reports, many early writers seemed to view financial panics the consequence of a contagion of pessimistic beliefs that become self-fulfilling prophesies. The classic Diamond and Dybvig (1983) model of banking was designed to formalize precisely this ancient view. Others, notably Gorton (1988), have questioned this view, arguing that the evidence suggests that bank panics are driven by fundamentals rather than psychology. Allen and Gale (1998) describe a version of the Diamond-Dybvig model where crisis events are driven entirely by fundamentals. A sort of hybrid of these two views is formalized by Goldstein and Pauzner (2005) who use a global games approach to make the probability of crisis dependent on some combination of beliefs and fundamentals.

The idea that fundamentals trigger bank crisis events is compelling, especially in light of the evidence provided by Gorton (1988). He shows that recessions Granger-cause bank panics in the U.S. National Banking Era 1863-1914. While there was no government deposit insurance or central bank during the National Banking Era, prominent bankers like J.P. Morgan would coordinate the formation of clearinghouses that issued certificates in lieu of cash. These clearinghouse certificates constituted claims against the combined assets of banks governed by the clearinghouse. In this way, clearinghouses performed a lender-of-last-resort function during bank panics. Indeed, Gorton (1988, table 1) reports that depositor losses during the seven panics in the years 1873-1914 were miniscule. One interpretation of this evidence is that such trivial losses are unlikely to have been possible if the underlying shock was fundamental, say, in the sense of Allen and Gale (1998). It may be that some fundamental bad news triggered a coordination failure. If so, then the Diamond and Dybvig (1983) approach remains relevant.[5]

Our model is also closely related to those in which money is introduced explicitly in theory. The list here includes Bryant (1980), Loewy (1991), Champ, Smith and Williamson (1996), Smith (2003), Jiang (2008), and Camous and Cooper (2014). With the exception of Camous and Cooper (2014), financial crises in this list of papers are triggered by changes

[4]See Lai (2002) for a useful survey of some more recent contributions.

[5]It would be a simple matter for us to include a leading indicator and compare bank runs precipated by rational news events versus irrational psychology events. There seems to be no clear consensus on which of these two views is correct–which probably means that both views possess an element of truth. For some interesting experimental work on the subject, see Arifovic, Jiang and Xu (2013), and the references cited within.

in economic fundamentals.

We use our model to the investigate the desirability and consequences of lender-of-last resort policies. Most of the theoretical literature on the subject is cast in real environments, including Diamond and Dybvig (1983). Papers like Bryant (1980) that examine the issue in the context of monetary economies resort to models where financial crises are driven by fundamentals. One appeal of fundamental crises is that equilibria are typically unique. A notable exception to this line of research includes a body of work that applies the concept of global games to determine uniqueness of runs driven by non-fundamentals. Rochet and Vives (2004) use such an approach to assess various issues related to lender-of-last-resort interventions, although theirs is a non-monetary model.

3 The environment

Time, denoted t, is discrete and the horizon is infinite, $t = 0, 1, 2, ..., \infty$. Each time period t is divided into three subperiods: the *morning*, *afternoon* and *evening*. There are two permanent types of agents, each of unit measure, which we label *investors* and *workers*.

Investors can produce morning output y_0 at utility cost $-y_0$. This output can be divided into consumer and capital goods, $y_0 = x + k$. Investors are subject to an idiosyncratic preference shock, realized at the end of the morning, which determines whether they prefer to consume early (in the afternoon) or later (in the evening). Let $0 < \pi < 1$ denote the probability that an investor desires early consumption c_1 (the investor is *impatient*). The investor desires late consumption c_2 (the investor is *patient*) with probability $1 - \pi$. (We assume that there is no aggregate uncertainty over investor types so that π also represents the fraction of investors who desire early consumption.) The utility payoffs associated with early and late consumption are given by $u(c_1)$ and $u(c_2)$, respectively, where $u'' < 0 < u'$ with $u'(0) = \infty$. Investors discount flow utility payoffs across periods with subjective discount factor $0 < \beta < 1$, so that investor preferences are given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[-x_t - k_t + \pi u(c_{1,t}) + (1 - \pi)u(c_{2,t})\right] \tag{1}$$

Workers have linear preferences for the morning and afternoon goods. In particular, workers wish to consume in the morning c_0 and have the ability to produce goods in the afternoon y_1. Goods produced in the afternoon can be stored into the evening at a unit gross rate of return. Workers share the same discount factor as investors, so that worker preferences are given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[c_{0,t} - y_{1,t}\right] \tag{2}$$

We now describe the investment technology available to investors. The rate of return on capital scrapped in the afternoon is $0 < \xi < 1$. The rate of return on capital that matures in the evening is $R > 1$.

To derive the properties of an efficient allocation, consider the problem of maximizing the *ex ante* welfare of investors, subject to delivering workers an expected utility payoff no less than $v = 0$. Since the problem is static, it may be written as

$$\max\left\{-x - k + \pi u(c_1) + (1-\pi)u(c_2)\right\} \tag{3}$$

subject to

$$\begin{aligned} x - \pi c_1 &\geq 0 && (4) \\ Rk + [x - \pi c_1] - (1-\pi)c_2 &\geq 0 && (5) \\ c_2 - c_1 &\geq 0 && (6) \end{aligned}$$

The condition $\xi < 1$ implies that workers can supply afternoon output more efficiently than investors who liquidate capital in the afternoon. As a consequence, efficiency dictates that early liquidation is never optimal and the problem above is formulated with this condition imposed. We have also imposed the resource constraint $x = c_0$. That is, the morning transfer of utility from investors to workers must add up (recall that there is an equal measure of investors and workers). Finally, (6) is the incentive constraint for patient agents: it ensures they do not misrepresent themselves. The incentive constraint for impatient agents, $c_1 \geq 0$ is trivially satisfied.

We first characterize an efficient allocation. To begin with, we can deduce that constraint (4) and (5) will bind tightly. The conditions that characterize the optimum are given by

$$\begin{aligned} u'\left(c_1^*\right) &= 1 \\ Ru'\left(c_2^*\right) &= 1 \end{aligned} \tag{7}$$

where, using (4) and (5), $c_1^* = x^*/\pi$ and $c_2^* = Rk^*/(1-\pi)$. Note that an efficient allocation satisfies, $x^*, k^* > 0$ and $0 < c_1^* < c_2^*$ and so the incentive compatibility constraint (6) holds.

4 A monetary economy

Before we discuss banking, it will be useful to introduce the frictions that imply a role for exchange media. To this end, assume that investors cannot commit to any promises they might make to workers, so that workers must be paid *quid-pro-quo* for output they produce in the afternoon. The lack of commitment implies a demand for an exchange medium, assumed here to take the form of a zero-interest-bearing government debt instrument (money), the total supply of which is denoted M_t at the beginning of date t. Assume that the initial money supply $M_0 > 0$ is owned entirely by workers. New money is created (destroyed) at the beginning of each morning at the constant rate $\mu \geq \beta$. New money $T_t = [M_t - M_{t-1}]$ is injected (withdrawn) as lump-sum transfers (taxes) bestowed (imposed) on workers.[6]

[6]While we permit any amount of deflation here in the range $\beta < \mu < 1$, there is the question of whether workers would be willing to pay the taxes necessary to finance any deflationary policy. Andolfatto (2013) addresses this issue, but we ignore it in what follows.

Trade of money-for-goods is assumed to take place in a sequence of competitive spot markets throughout the morning and afternoon, at prices p_t^m and p_t^a, respectively.[7] We anticipate a sequence of spot trades that consist of investors selling their morning production for money and using the cash proceeds to purchase output in the afternoon.

Consider a worker who enters the morning with m_{t-1} units of money, supplemented with the transfer T_t. For every unit of output a worker sells in the afternoon, he receives p_t^a units of money, which he could potentially sell for $1/p_{t+1}^m$ units of the morning good in the following morning period. Since his preferences in the afternoon and the following morning are linear, the following condition has to hold:

$$1/p_t^a = \beta/p_{t+1}^m \tag{8}$$

In a stationary monetary equilibrium, all nominal prices grow at the same rate μ as the aggregate stock of money. Thus, $p_{t+1}^m/p_t^m = p_{t+1}^a/p_t^a = \mu$ and so from (8):

$$p_t^m/p_t^a = \beta/\mu \tag{9}$$

Since $\mu \geq \beta$, we anticipate that investors will enter the morning with zero money balances, accumulate money in the morning and spend all their money in the afternoon.[8] Thus, the quantity

$$m_t = p_t^m x \tag{10}$$

represents both the nominal value of cash acquired by an investor in the morning and held into the afternoon.

There is a secondary capital market in the afternoon that permit patient and impatient investors to trade afternoon output for claims to evening output. Let η denote the probability that this market is closed. We interpret the possibility of closed market as a market freeze. It is related to a sunspot equilibrium. Assume that in period t a signal s_t is drawn from the uniform distribution on $[0, 1]$. The draw is unrelated to any other variables in the economy and is independent across periods (Ennis and Keister, 2003). The realization of the sunspot signal s_t is observed by all investors. Investors follow the decision rule "do not enter the market if $s_t \leq \eta$; otherwise enter." As in any standard model with sunspots, the sunspot signal s_t coordinates the behavior of agents on one of the equilibria. If an investors believes that the market is closed, a best response is to not enter since there is nothing to trade. If he believes that the market is open, the best response is to enter and trade.

Assume that the market is open. Denote p_t^k the price of capital in the afternoon and $\rho = p_t^k/p_t^a$ the relative price. Furthermore, denote k^P and k^I the capital bought and the capital sold by a patient and impatient investor, respectively. In this set up, the unsold capital of an impatient investor is liquidated. If the investor turns out to be impatient, he faces the expenditure constraint

$$c_1 \leq \frac{m_t}{p_t^a} + \rho k^I + \xi\left(k - k^I\right) \tag{11}$$

[7] There is no spot market in the evening because workers cannot produce and do not desire consumption in the evening.

[8] Note that the money investors spend in the afternoon is used to purchase output that, by assumption, can be stored into the evening.

If the investor turns out to be patient, he carries m_t/p_t^a units of afternoon goods into the evening so that consumption is constrained by

$$c_2 \leq \frac{m_t}{p_t^a} - \rho k^P + R\left(k + k^P\right) \tag{12}$$

Combine (10) and (11) set to equality, together with (9), to form $c_1 = (\beta/\mu)x + \rho k^I + \xi\left(k - k^I\right)$. Similarly, we can transform (57) into $c_2 = (\beta/\mu)x - \rho k^P + R\left(k + k^P\right)$.

If the market is not open, the capital investment associated with an impatient investor needs to be scrapped. If the investor turns out to be impatient, he faces the expenditure constraint

$$\hat{c}_1 \leq \frac{m_t}{p_t^a} + \xi k \tag{13}$$

If the investor turns out to be patient, he carries m_t/p_t^a units of afternoon goods into the evening so that consumption is constrained by

$$\hat{c}_2 \leq \frac{m_t}{p_t^a} + Rk \tag{14}$$

Combine (10) and (13) set to equality, together with (9), to form $\hat{c}_1 = (\beta/\mu)x + \xi k$. Similarly, we can transform (59) into $\hat{c}_2 = (\beta/\mu)x + Rk$.

The investor's objective function satisfies

$$\max -x - k + (1-\eta)\left\{\pi u\left(c_1\right) + (1-\pi)u(c_2)\right\} + \eta\left\{\pi u\left(\hat{c}_1\right) + (1-\pi)u(\hat{c}_2)\right\} \tag{15}$$

subject to

$$\begin{aligned} k - k^I &\geq 0 \\ (\beta/\mu)x - \rho k^P &\geq 0 \end{aligned}$$

The first constraint means that the impatient investor cannot sell more capital than he holds, whereas the second constraint says that the patient investor can not spend more cash than he holds. Denote the Lagrange multipliers for these constraint λ_k^I and λ_k^P, respectively. When $(1-\eta)\left[\pi\rho + (1-\pi)R\right] + \eta\xi > \beta/\mu$, the expected return on capital is higher than the return on money and thus, investors will hold some capital.

An investor's desired portfolio choice is characterized by

$$\begin{aligned} (1-\eta)[\pi u'(c_1) + (1-\pi)u'(c_2)] + \eta[\pi u'(\hat{c}_1) + (1-\pi)u'(\hat{c}_2)] + \lambda_k^P &= \mu/\beta & (16) \\ (1-\eta)[\pi\xi u'(c_1) + (1-\pi)Ru'(c_2)] + \eta[\pi\xi u'(\hat{c}_1) + (1-\pi)Ru'(\hat{c}_2)] + \lambda_k^I &= 1 & (17) \\ (1-\eta)\pi u'(c_1)(\rho - \xi) - \lambda_k^I &= 0 & (18) \\ u'(c_2)(R-\rho) - \lambda_k^P \rho &= 0 & (19) \end{aligned}$$

We have the endogenous variables $(x, k, k^I, k^P, \lambda_k^I, \lambda_k^P, \rho)$. Equations (61)–(64), the Kuhn-Tucker conditions $\lambda_k^I(k - k^I) = 0$ and $\lambda_k^P[(\beta/\mu)x - \rho k^P] = 0$, and the market clearing condition

$$\pi k^I = (1-\pi)\, k^P$$

solve for these.

Lemma 1 *In a monetary equilibrium, $\xi \leq \rho \leq R$ (wlog when $\eta = 1$).*

Proof. Given $\lambda_k^I \geq 0$ and $\lambda_k^P \geq 0$, conditions (63) and (64) imply $(1-\eta)\pi u'(c_1)(\rho - \xi) \geq 0$ and $u'(c_2)(R - \rho) \geq 0$. In a monetary equilibrium, given Inada conditions, $c_1, c_2 > 0$. Thus, $\rho - \xi > 0$ (wlog when $\eta = 1$) and $R - \rho > 0$. ■

Use (63) and (64) to solve for λ_k^I and λ_k^P. Conditions (61) and (62) imply:

$$
\begin{aligned}
(1-\eta)[\pi u'(c_1) + (1-\pi)(R/\rho)u'(c_2)] + \eta[\pi u'(\hat{c}_1) + (1-\pi)u'(\hat{c}_2)] &= \mu/\beta \qquad (20)\\
(1-\eta)[\pi \rho u'(c_1) + (1-\pi)Ru'(c_2)] + \eta[\pi \xi u'(\hat{c}_1) + (1-\pi)Ru'(\hat{c}_2)] &= 1 \qquad (21)
\end{aligned}
$$

If $\xi < \rho < R$, then $\lambda_k^I, \lambda_k^P > 0$ and thus, $k = k^I$ and $(\beta/\mu)x = \rho k^p$. Given the market clearing condition $\pi k^I = (1-\pi)k^p$ we obtain

$$\rho = \frac{\beta}{\mu}\frac{x}{\pi}\frac{(1-\pi)}{k} \qquad (22)$$

and $c_1 = (\beta/\mu)(x/\pi)$, $c_2 = Rk/(1-\pi)$, $\hat{c}_1 = (\beta/\mu)x + \xi k$ and $\hat{c}_2 = (\beta/\mu)x + Rk$.

The initial portfolio (x, k) constitutes part of a stationary monetary equilibrium. To recover equilibrium nominal values, impose the equilibrium condition $m_t = M_t$. From condition (10) we then have an expression for the morning price-level, $p_t^m = M_t/x$. Condition (9) then implies $p_t^a = (\mu/\beta)\, p_t^m$, and so on.

In this monetary economy, investors are motivated to accumulate both money and capital in the morning. They acquire money by selling consumer goods to the workers and they accumulate capital directly with their own effort. In the afternoon, an impatient investor spends his money for afternoon goods (supplied to him by workers) and sells his capital. A patient investor spends all his money for capital in the afternoon and in the evening, consumes the return to the maturing investment.

Proposition 2 *The first best cannot be implemented in a monetary equilibrium.*

Proof. At the first best, we require $c_1 = \hat{c}_1 = c_1^*$ and $c_2 = \hat{c}_2 = c_2^*$ if $\eta > 0$ or $c_1 = c_1^*$ and $c_2 = c_2^*$ if $\eta = 0$. Impose conditions (7) into (61)–(64). We get

$$
\begin{aligned}
(1-\pi)(1/R) + \lambda_k^P &= (1-\pi)(\mu/\beta)\\
\pi\xi(\mu/\beta) + \lambda_k^I &= \pi\\
(1-\eta)\pi(\mu/\beta)(\rho - \xi) - \lambda_k^I &= 0\\
(1 - \rho/R) - \lambda_k^P \rho &= 0
\end{aligned}
$$

Since $\mu \geq \beta$ and $R > 1$, the first condition above cannot be satisfied for $\lambda_k^P = 0$. In any monetary equilibrium, $\beta/\mu \geq \xi$ (otherwise, no agent holds money). Suppose $\beta/\mu > \xi$, then the second condition above cannot be satisfied for $\lambda_p^I = 0$. Thus, $\lambda_k^I, \lambda_k^P > 0$ and so $\xi < \rho < R$; In particular, $\rho^* = (\beta/\mu)(x^*/\pi)(1-\pi)/k^*$. We further get

$$
\begin{aligned}
\pi(1/R) + (1-\pi)(\mu/\beta) &= 1/\rho \qquad (23)\\
(\mu/\beta)[\eta\xi + (1-\eta)\rho] &= 1 \qquad (24)
\end{aligned}
$$

Note that $c_1 = \hat{c}_1$ implies $(\beta/\mu)x/\pi = (\beta/\mu)x + \xi k$, which implies $\rho = \xi$, a contradiction. Thus, we cannot have $\eta > 0$. Intuitively, scrapping is inefficient, so it cannot happen at the first best. With $\eta = 0$, (24) implies $\rho = \beta/\mu$ and so (23) implies $1/R = \mu/\beta$, a contradiction.

Now suppose $\beta/\mu = \xi$. Then, $\lambda_k^I = 0$ and so, $\rho = \xi$ and $\lambda_k^P = 1/\xi - 1/R > 0$. Next, we get the analog of (23)

$$\pi(1/R) + (1-\pi)(\mu/\beta) = 1/\xi$$

Since $\mu/\beta = 1/\xi$, the equation above implies $1/R = 1/\xi$, a contradiction. ■

The proposition above shows that we cannot implement the first best allocation in a monetary equilibrium. This is true even for the case when the monetary authority is running the Friedman rule, $\mu = \beta$ and the secondary market is always open $\eta = 0$.

One way to overcome these inefficiencies is to implement a complete contingent-claims market. Doing so is straightforward if investor types were observable and contractible. But since investor type is private information, such claims must be made contingent on incentive-compatible reports. One solution to the problem of sharing risk with private information entails the creation of a bank that funds its operations with demandable debt, as in Diamond and Dybvig (1983). But unlike the Diamond-Dybvig model, deposit liabilities here are optimally designed to be redeemable in fiat money (instead of direct claims on goods). This is a distinction we believe to be important in a world of nominal bank debt redeemable in an object under the direct control of a central bank.

5 A Diamond-Dybvig bank

We want to take one more preliminary step before getting into our full model. To this point, we have described the environment and the nature of monetary exchange. We want to now explain how banking fits into our model under the provisional assumption that bank runs are not expected to occur.

To exploit the gains associated with risk-sharing, we assume that investors coalesce to form a Diamond-Dybvig (1983) style bank. The banking arrangement works as follows. Investors work in the morning, creating y_0 units of output (a mix of consumer and capital goods) in exchange for bank-money (the bank funds this purchase with its own liabilities). The bank sells $x \leq y_0$ units of this output, in the form of consumer goods, for cash $m_t = p_t^m x$. The remaining output $k = y_0 - x$ is formed into capital. Bank-money constitutes a claim against the bank's assets $x + k$. The bank's deposit liabilities are made redeemable for cash because it is known and expected that impatient investors will want to cash out. The redemption option is made demandable because preferences are private information. Deposit liabilities not redeemed in the afternoon constitute *pro rata* claims against the output generated by the maturing investment in the evening.

5.1 Optimal risk-sharing

We begin by describing the optimal risk-sharing arrangement in the monetary economy ignoring–for the moment–the possibility of multiple equilibria. We conjecture (and later verify) that if $\mu > \beta$, a bank will enter each morning with zero money balances (and wlog if $\mu = \beta$). Thus, with m_t dollars acquired in the morning, afternoon consumption is limited by $\pi p_t^a c_1 \leq m_t$. Combining $m_t = p_t^m x$ and condition (9) with this latter restriction yields the constraint

$$(\beta/\mu)\, x \geq \pi c_1. \tag{25}$$

Thus, the bank holds cash reserves in sufficient quantities to meet expected redemptions. Any cash left over following afternoon redemption activity $[(\beta/\mu)\, x - \pi c_1]$ is used to purchase afternoon output and then stored to finance consumption in the evening.[9] The bank therefore faces an evening state-contingent budget constraint

$$Rk + [(\beta/\mu)\, x - \pi c_1] \geq (1-\pi)c_2. \tag{26}$$

The bank's choice problem is to maximize

$$\max \left\{ -x - k + \pi u(c_1) + (1-\pi)u(c_2) \right\}. \tag{27}$$

subject to (25), (68) and the incentive constraint (6).

The solution to this problem implies constraints (25) and (68) bind, whereas constraint (6) remains slack. For a given monetary policy μ, the equilibrium consumption allocation is characterized by

$$u'(c_1) = \mu/\beta \tag{28}$$
$$Ru'(c_2(R)) = 1 \tag{29}$$

with $c_1 = (\beta/\mu)(x/\pi)$ and $c_2 = Rk/(1-\pi)$. Note that (28) and (29) imply $0 < c_1 < c_2$ and so the incentive compatibility constraints are satisfied. Also note that by (7), (29) implies $c_2 = c_2^*$ and $k = k^*$.

Proposition 3 *In a monetary economy with Diamond-Dybvig style banks, the Friedman rule $\mu = \beta$ implements the first best.*

Proof. Plugging $\mu = \beta$ into (28) implies $u'(c_1) = 1$ and so by (7) $c_1 = c_1^*$ and hence, $x = x^*$. ■

5.2 Unanticipated bank run equilibrium

In this section, we describe our version of the Diamond-Dybvig "bank run" equilibrium. The thought experiment here takes the allocation above, derived under the assumption

[9] One might think that the bank might alternatively hold the cash and carry it over into the next period. Doing so, however, is suboptimal if $\mu < \beta$. In this case, the bank will only accumulate cash if it intends to disburse it or spend it.

that no bank run is possible, and asks whether a bank run equilibrium exists. (Later, we will derive the optimal banking arrangement when the bank recognizes the existence of run equilibria.)

We restrict attention to pure strategy equilibria so that patient investors either "show up" in the afternoon exercising their redemption options *en masse*, or they wait for the investment to mature in the evening. Although we could, we do not impose a sequential service constraint in the afternoon.[10] Therefore, if patient investors "run" the bank in the afternoon–and if the bank is unable to honor its short-term obligations–the bank is programmed to pay out a *pro rata* share of *all* available resources to *all* investors who appear in the afternoon.[11] Note that the spirit of sequential service is preserved here in the sense that patient investors will not be serviced if they fail to arrive early (in the afternoon).[12]

The bank has promised to pay (in real terms) c_1 units of output to any investor wanting to make a withdrawal in the afternoon. If only impatient investors exercise the early redemption option, then (25) guarantees that the bank's promise can be honored. But if patient investors arrive *en masse* to make withdrawals, then honoring all withdrawal requests requires $c_1 \leq (\beta/\mu)x$. However, since constraint (25) binds, $c_1 = (\beta/\mu)(x/\pi) > (\beta/\mu)x$ and hence, in order to meet its short-term obligations, the bank needs to sell/liquidate capital. As for the monetary economy, ρ is the competitive price of capital and k^B denotes the quantity of capital sold. Accordingly, $k - k^B$ is the quantity of capital that is scrapped. The amount of purchasing power the bank can dispense through selling/liquidation is given by

$$\tilde{c}_1 = (\beta/\mu)x + \rho k^B + \xi\left(k - k^B\right) \tag{30}$$

Evidently, if $\rho \geq \xi$ a weakly dominant strategy is to sell.[13] Note that it is possible here that $\tilde{c}_1 > c_1$, in which case the bank need only to sell/liquidate some of its capital to honor its redemption promise. For the purpose of explaining the basic idea, we can assume here that ρ and ξ are sufficiently close to zero so that $\tilde{c}_1 \leq c_1$ in which case any delay on the part of investors assures that they will receive nothing in the evening in the event of a run.[14] In this case, it is immediately evident that a bank run equilibrium exists.

In this case, when there is a run, the bank sells $k^B = k$ units of capital at price ρ and the $(1-\pi)$ patient investors buy k^P units each at the same price. Accordingly, the market clearing condition is

$$k = (1-\pi)\,k^P$$

[10] Here, we follow Allen and Gale (1998).

[11] We assume that the bank cannot credibly commit to suspend redemptions in the event of a run. As is well-known, a credible threat to suspend early redemptions once reserves are depleted will prevent bank run equilibria in this environment. Andolfatto, Nosal and Sultanum (2014) demonstrate that bank runs can be prevented in a wide class of environments under appropriately designed indirect mechanisms. Their result suggests that some form of market incompleteness is necessary to admit the possibility of bank run equilibria.

[12] Further subjecting investors to sequential service in the afternoon induces a lottery over consumption, which would complicate the model in an interesting way, but is otherwise unrelated to the existence of multiple equilibria.

[13] Later we show that $\rho > \xi$ in any equilibrium.

[14] Later we show under which condition $\tilde{c}_1 = c_1$.

Both quantities are chosen optimally by the bank in order to maximize the lifetime discounted utility of the representative investor as shown further below. When doing so, the bank takes the price of capital ρ as given. Note, further, that the selling of capital when there is a run also implies that evening consumption changes. Thus, consumption quantities satisfy

$$\begin{aligned} \tilde{c}_1 &= (\beta/\mu)x + \rho k && (31) \\ \tilde{c}_2(R) &= (\beta/\mu)x + \rho k + (R-\rho)k^P && (32) \end{aligned}$$

The analysis above assumes that the event triggering a bank run (e.g., a "sunspot") is completely unanticipated. Below, we assume that sunspots occur with a known probability θ. If this is so, then banks can be expected to alter the size and composition of their balance sheet depending on the degree of "inherent fragility" characterizing their social environment, as indexed by the parameter θ. Doing so will permit us to ask how bank behavior changes with θ and whether banking is even desirable in high θ environments.

As for the monetary economy, we assume that the secondary capital market is only open with probability η. If it is closed, the DD-bank liquidates the capital and consumption for all investors satisfies

$$\hat{c}_1 = (\beta/\mu)x + \xi k \tag{33}$$

6 Money and banking in fragile societies

Assume an exogenous stochastic process $\{s_t\}_{t=0}^{\infty}$ where $s_t \in \{0,1\}$ is a "sunspot" variable realized at the beginning of the afternoon in period t. Let $\theta \equiv \Pr[s_t = 1]$ for any t (it is straightforward to permit time-dependence, but we stick to the *i.i.d.* case below). We further assume that impatient investors demand early redemptions *en masse* when $s_t = 1$ and otherwise do not misrepresent themselves when $s_t = 0$. As explained above, such behavior constitutes an equilibrium for any $0 \leq \theta \leq 1$. Because θ reflects the propensity of bank runs in an economy, we think of θ as indexing the degree of " inherent fragility" in an economy–that is, the propensity for a given society to coordinate on a suboptimal outcome. Our goal here is to examine how a bank may alter the size and composition of its asset portfolio when it understands that the economy it is working in is prone to coordination failure.

In the event of a run, we assume parameters such that full liquidation/selling of the bank's assets is insufficient to meet its short-term obligations, so that $\tilde{c}_1, \hat{c}_1 \leq c_1$ in any equilibrium (where $\tilde{c}_1$ is given by (31) and $\hat{c}_1$ by (33)). Also, because $\mu > \beta$, we anticipate that the bank will not want to carry cash into the following period. In this case, the bank's choice problem can be stated as follows

$$W^B(\theta,\mu) \equiv \max \left\{ \begin{array}{l} -x - k + (1-\theta)\left[\pi u(c_1) + (1-\pi)u(c_2)\right] \\ +(1-\eta)\,\theta[\pi u(\tilde{c}_1) + (1-\pi)u(\tilde{c}_2)] + \eta\theta u\left(\hat{c}_1\right) \end{array} \right\} \tag{34}$$

subject to (31)–(33), the incentive constraint (6), and

$$
\begin{aligned}
(\beta/\mu)\, x &\geq \pi c_1 && (35)\\
Rk + [(\beta/\mu)\, x - \pi c_1] &\geq (1-\pi)c_2 && (36)\\
(\beta/\mu)x + \rho(k - k^P) &\geq 0 && (37)
\end{aligned}
$$

Constraint (37) specifies that, in the event of a run, patient agents cannot buy more capital, valued ρk^p, than the cash they withdrawn from the bank, $(\beta/\mu)x + \rho k$.

It is easy to show that as long as $c_1 < c_2$, the cash reserve constraint (77) binds. With Lagrange multiplier λ_k^P on (37), the solution to the bank's problem is characterized by the following set of conditions

$$
\begin{aligned}
(1-\theta)u'(c_1) + (1-\eta)\theta[\pi u'(\tilde{c}_1) + (1-\pi)u'(\tilde{c}_2)] + \eta\theta u'(\hat{c}_1) + \lambda_k^P &= \mu/\beta && (38)\\
(1-\theta)Ru'(c_2) + (1-\eta)\theta\rho[\pi u'(\tilde{c}_1) + (1-\pi)u'(\tilde{c}_2)] + \eta\theta\xi u'(\hat{c}_1) + \lambda_k^P\rho &= 1 && (39)\\
(1-\eta)\theta(1-\pi)u'(\tilde{c}_2)(R-\rho) - \lambda_k^P\rho &= 0 && (40)
\end{aligned}
$$

These first-order conditions plus the complementary slackness condition, $\lambda_k^P[(\beta/\mu)x + \rho(k - k^P)]$ and the market clearing condition, $k = (1-\pi)k^P$, solve for $\left(x, k, \rho, k^P, \lambda_k^P\right)$.

Condition (40) implies $\lambda_k^P = (1-\eta)\theta(1-\pi)u'(\tilde{c}_2)(R/\rho - 1) \geq 0$ and so, $\rho \leq R$. Plugging into (38) and (39) we obtain

$$
\begin{aligned}
(1-\theta)u'(c_1) + (1-\eta)\theta[\pi u'(\tilde{c}_1) + (1-\pi)(R/\rho)u'(\tilde{c}_2)] + \eta\theta u'(\hat{c}_1) &= \mu/\beta && (41)\\
(1-\theta)Ru'(c_2) + (1-\eta)\theta[\pi\rho u'(\tilde{c}_1) + (1-\pi)Ru'(\tilde{c}_2)] + \eta\theta\xi u'(\hat{c}_1) &= 1 && (42)
\end{aligned}
$$

If $\xi < \rho < R$ then $\lambda_k^P > 0$ and so, $(\beta/\mu)x = \rho(k^P - k)$. Using the market clearing condition, $k = (1-\pi)k^P$ we obtain

$$
\rho = \frac{\beta}{\mu}\frac{x}{\pi}\frac{(1-\pi)}{k} \qquad (43)
$$

and so, $c_1 = (\beta/\mu)(x/\pi)$, $c_2 = Rk/(1-\pi)$, $\tilde{c}_1 = (\beta/\mu)x + \rho k$, $\tilde{c}_2 = (\beta/\mu)x + [k/(1-\pi)](R - \pi\rho)$ and $\hat{c}_1 = (\beta/\mu)x + \xi k$.

Proposition 4 *In a monetary economy with Diamond-Dybvig style banks and sunspots, banks meet all their afternoon obligations when there is a run and the secondary market is open. Specifically,* $\tilde{c}_1 = c_1$ *and* $\tilde{c}_2 = c_2$.

Proof. Using (43), we obtain $\tilde{c}_1 = (\beta/\mu)(x/\pi) = c_1$ and $\tilde{c}_2 = Rk/(1-\pi) = c_2$. ■

Therefore, the equilibrium in a monetary economy with a fractional reserve banking system subject to runs and market freezes is characterized by:

$$
\begin{aligned}
(1-\theta)u'(c_1) + (1-\eta)\theta[\pi u'(c_1) + (1-\pi)(R/\rho)u'(c_2)] + \eta\theta u'(\hat{c}_1) &= \mu/\beta && (44)\\
(1-\theta)Ru'(c_2) + (1-\eta)\theta[\pi\rho u'(c_1) + (1-\pi)Ru'(c_2)] + \eta\theta\xi u'(\hat{c}_1) &= 1 && (45)
\end{aligned}
$$

with ρ given by (43), and $c_1 = (\beta/\mu)(x/\pi)$, $c_2 = Rk/(1-\pi)$ and $\hat{c}_1 = (\beta/\mu)x + \xi k$.

Note that a run is still an equilibrium even when $\eta = 0$ as the bank needs to sell all its capital in order to meet all its afternoon obligations. What happens when the secondary market is always open? When $\eta = 0$, (44) and (45) simplify to

$$\begin{aligned}(1-\theta)u'(c_1) + \theta[\pi u'(c_1) + (1-\pi)(R/\rho)u'(c_2)] &= \mu/\beta \\ (1-\theta)Ru'(c_2) + \theta[\pi\rho u'(c_1) + (1-\pi)Ru'(c_2)] &= 1\end{aligned}$$

When $\eta = 0$, as $\theta \to 0$, the economy with sunspots converges to the Diamond-Dybvig case without sunspots; as $\theta \to 1$, the economy with sunspots converges to the monetary equilibrium. For interior values of θ, the economy with sunspots is a convex combination of the other two cases. As such, it cannot achieve the first best, even when running the Friedman rule.

Proposition 5 *In a monetary economy with Diamond-Dybvig style banks and sunspots, the Friedman rule does not implement the first best.*

6.1 Narrow banking

The benefits of risk-sharing that banking entails are likely to be weighed against the cost that this financially fragile structure imposes (here, in the way of inefficient liquidation). As we showed above, the welfare benefit associated with banking declines monotonically in θ, the probably of coordination failure. If the frequency of coordination failure is sufficiently high, it could turn out that fractional reserve banking is dominated by an alternative regime. The alternative regime we consider here is a narrow banking regime (100% reserve requirement).

In our model, liquidation is assumed to occur when the bank cannot honor its short-term obligations. In a narrow banking regime, banks are restricted to making short-term obligations that they can honor in *every* state of the world. In other words, demandable debt has to be fully backed with cash reserves. There is an obvious cost to this restriction, since the bank must hold additional idle cash at the expense of more profitable investments. But on the other hand, it enjoys a great benefit in that the structure is run-proof. In other words, the allocation attained under a narrow banking regime is independent of θ. Of course, this may matter a great deal for intermediaries operating in high θ economies.

Since patient investors never have an incentive to run a narrow bank, the narrow bank's objective need not consider the effect of θ. Similarly, whether the afternoon market for capital is open is not relevant. That is, the bank regime objective (76) now reduces to

$$W^N(\mu) \equiv \max\{-x - k + \pi u(c_1) + (1-\pi)u(c_2)\} \tag{46}$$

The constraints are now given by

$$\begin{aligned}(\beta/\mu)\,x &\geq c_1 && (47)\\ Rk + [(\beta/\mu)\,x - \pi c_1] &\geq (1-\pi)c_2 && (48)\end{aligned}$$

Notice that the cash reserve constraint (47) is significantly more stringent than (77) and will thus, also bind. Therefore, we get $c_1 = (\beta/\mu)x$ and from (48), $c_2 = Rk/(1-\pi) + (\beta/\mu)x$.

The first-order conditions imply:

$$\begin{aligned} \pi u'(c_1) + (1-\pi)u'(c_2) &= \mu/\beta \\ Ru'(c_2) &= 1 \end{aligned}$$

Note that $c_2 = c_2^*$, as in the case with the Diamond-Dybvig type bank without runs. However, c_1 is lower than in that case, as the narrow bank is forced to hold excess cash to avoid runs. In addition, capital investment is lower as part of patient agents' consumption is financed with cash.

The following two results follows immediately.

Proposition 6 ***Effect of inflation on narrow banking asset composition:*** *$dx/d\mu < x/\mu$ and $dk/d\mu > 0$.*

In the narrow bank case, inflation increases capital investment; the effect on real money balances is, however, ambiguous. All we can say in this latter case is that the elasticity of real money demand with respect to inflation is less than unity $(\mu/x)dx/d\mu < 1$. In terms of how these parameters affect the franchise value of narrow banks, we offer the following proposition.

Proposition 7 *The economic benefit of narrow banking $W^N(\mu)$ is independent of the level of social fragility θ and strictly decreasing in the rate of inflation μ.*

We conjecture that, when $\eta = 1$ (no afternoon capital market), investors prefer to form fractional reserve banks in low θ economies and narrow banks in high θ economies. The intuition is straightforward, high θ economies are subject to frequent inefficient liquidation events. In the limit, as $\theta \to 1$, a liquidation event occurs in every period. In the opposite case, as $\theta \to 0$, we have already established that the Diamond and Dybvig fractional reserve bank implements an efficient allocation (under the Friedman rule). Consequently, there must exist a $\hat{\theta}$ such that investors are just indifferent between a fractional and 100% reserve banking system.

Proposition 8 *For a wide range of parameter values, there exists a unique $\hat{\theta}(\mu) \in (0,1)$ that satisfies $W^B(\hat{\theta}, \mu) \equiv W^N(\mu)$. Moreover, $\hat{\theta}(\mu)$ is strictly increasing in μ.*

Proof. At the moment, proof of existence is given by example. But intuitively, consider the following argument. It is easy to see that W^B in (76) is strictly decreasing in θ. As $\theta \to 0$, the fractional reserve bank allocation approaches the efficient allocation (7), at least for sufficiently small inflation rate, so that $\lim_{\theta \to 0} W^B(\theta, \mu) > W^N(\mu)$. As $\theta \to 1$, the fractional reserve bank allocation is a disaster–inefficient liquidation events occur at high frequency, so that $\lim_{\theta \to 0} W^B(\theta, \mu) < W^N(\mu)$. The narrow bank allocation is independent of θ. By the theorem of the maximum, the value functions W^B, W^N are continuous. The strict monotonicity of W^B in θ establishes the existence of the cut off value $\hat{\theta}(\mu)$.

By the envelope theorem, one can show that both $W^B(\theta,\mu)$ and $W^N(\mu)$ are strictly decreasing functions of the inflation rate μ. At first blush then, it seems difficult to establish analytically the properties of the function $\hat{\theta}(\mu)$. However, one can demonstrate analytically that the sign of $\hat{\theta}'(\mu)$ is the same as the sign of $x^N - x^B$, that is, the difference in the demand for real cash balances across the two regimes. As it turns out, $x^N > x^B$, that is, perhaps not surprisingly, the demand for real cash balances is higher in the narrow bank regime. Because this is the case, an increase in the inflation rate penalizes the narrow bank regime relatively more than it does the fractional reserve bank regime. That is, recall that inflation is a tax on non-interest bearing money. ■

6.2 Inflation and financial fragility

The parameter θ in our theoretical framework can be thought of as indexing the inherent fragility of an economy–its propensity for coordination failure. The structure of an economy's banking system, however, is perhaps better thought of as a manufactured fragility. It is not unreasonable to suppose that the structure of a banking system might accommodate itself to the inherent properties of the environment it operates in.

The previous proposition has the following interesting implication: we should not expect to obverse fragile banking systems in inherently fragile economies. Formally, we would expect narrow banking structures to emerge in high θ economies ($\theta > \hat{\theta}(\mu)$). The fragile fractional reserve banking structure is more likely to emerge in low θ ($\theta < \hat{\theta}(\mu)$) environments–economies that are characterized by a greater degree of inherent social stability.

Our model therefore provides a theory that explains the probability of financial crisis. Among other things, our theory suggests that the probability of crisis is related to the inflation rate (induced by financing primary budget surpluses with money creation). By the proposition above, high inflation economies are associated with a higher $\hat{\theta}(\mu)$, that is, a higher tolerance for fractional reserve banking systems–and hence, a greater likelihood of financial crisis. High inflation environments make stable banking systems expensive to operate–the inflation tax punishes banks with large non-interest-bearing cash reserves. Consider two inflation rates $\mu_l < \mu_h$ and consider investors operating in an economy θ where $\hat{\theta}(\mu_l) < \theta < \hat{\theta}(\mu_h)$. These investors will prefer a stable narrow banking system in the low inflation regime, but will convert to a fractional reserve banking system in a high inflation regime.

6.3 Wall street vs main street

It is of some interest to note that the benefits that accrue to investors through their choice of banking regime are not always shared with other interests–in our model, the class of workers. Consider, for example, an economy for which $\theta = \hat{\theta}(\mu)$. It turns out that while investors are indifferent with respect to bank regime, workers would strictly prefer a narrow bank regime.

Workers' welfare is straightforward to compute. In any of the cases considered above, a worker exchanges his money holdings in the morning for x units of transferable utility; in the afternoon, he works the equivalent of $(\beta/\mu)x$. Thus, a worker's flow utility is equal to $x(1-\beta/\mu) \geq 0$. For both fractional and narrow banking, our numerical simulations suggest that workers' welfare is increasing in δ and μ; in terms of θ, welfare is increasing in the fractional reserve case and constant with narrow banking.

Given that fractional reserve banking economizes on real cash balances x, investors and workers disagree on the desirability of this institution. Typically, investors prefer fractional banking unless θ is too large (when the regime is too fragile), whereas workers prefer narrow banking. It is interesting to note, however, that if θ is large enough, workers may prefer fractional banking while investors would prefer narrow banking. This is a scenario where real cash balances are higher under fractional reserve than narrow banking.

6.4 Numerical Analysis

To investigate the properties of the allocation, we turn to a numerical example. Let $u(c) = (1-\sigma)^{-1}\left[c^{1-\sigma}-1\right]$. Benchmark parameter values are set as follows:

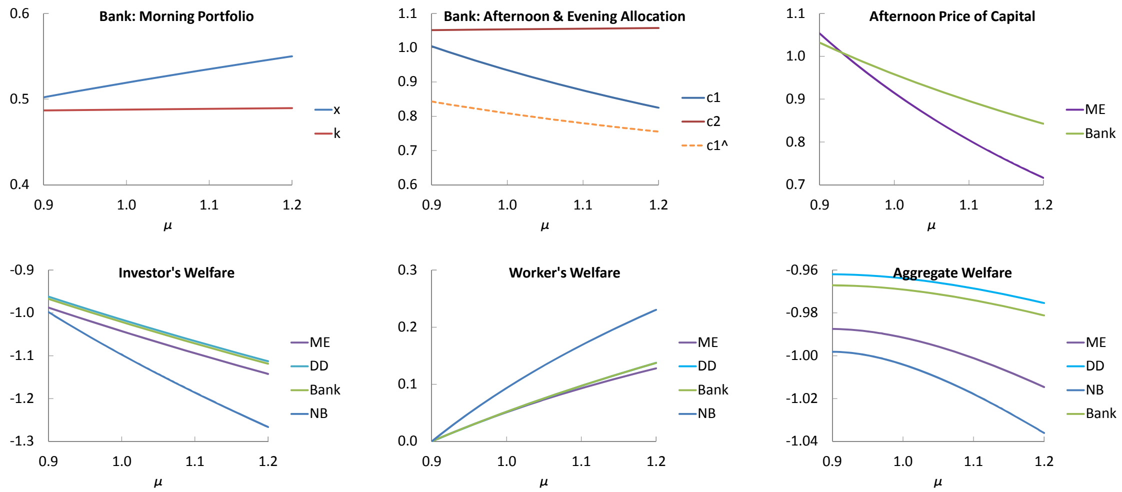
$$\beta = 0.90 \quad \mu = 1.05 \quad \sigma = 1.50 \quad R = 1.08 \quad \theta = 0.10 \quad \eta = 0.25 \quad \xi = 0.70 \quad \pi = 0.50 \tag{49}$$

In what follows, we report results for how the allocation varies with monetary policy (inflation rate financed by lump-sum transfers to workers) (μ), the frequency of the sunspot (θ) and the frequency the afternoon secondary market for capital is open (η). See Figures 1, 2 and 3, respectively. Each of these figures show: the allocation for the case of fractional reserve banking with sunspots; the afternoon price of capital in the secondary market (when applicable); and investors, workers and overall welfare for all cases considered.

Figure 1 shows the effects of increasing inflation. In a fractional reserve banking system, as expected, afternoon consumption is decreasing in inflation, while evening consumption is increasing, although slightly. In the event of a concurrent run and market freeze, the afternoon allocation is decreasing in inflation. In terms of the morning portfolio, there is a minor Tobin effect, which increases capital with inflation (with $\theta = \eta = 0$ capital would not change with inflation). More surprisingly, the demand for money increases with inflation. This result follows from our assumption on u; the opposite effect would obtain if the curvature on the utility function was lower than log. In the secondary market, the price of capital is decreasing in inflation. In terms of welfare, for our benchmark parameters, the possibility of runs and market freezes has a minor impact. Still, a fractional banking system dominates the monetary equilibrium without banks. A narrow banking system lowers the welfare of investors but increases the welfare of workers. These loses and gains get more pronounced as inflation increases.

Figure 2 shows the effects of increasing the frequency of runs. In a fractional reserve banking system, all allocations are increasing in θ. Here the interplay of runs with the possibility of a market freeze and the rate of inflation is critical. For example, a different parameter combination, would show money holdings and afternoon consumption decreasing in θ. In all cases we looked at, however, the price of capital in the secondary market decreases with more frequent runs. The welfare implications are interesting: as runs become more frequent, investors would rather not form banks. Workers, on the other hand,

Figure 1: Inflation rate

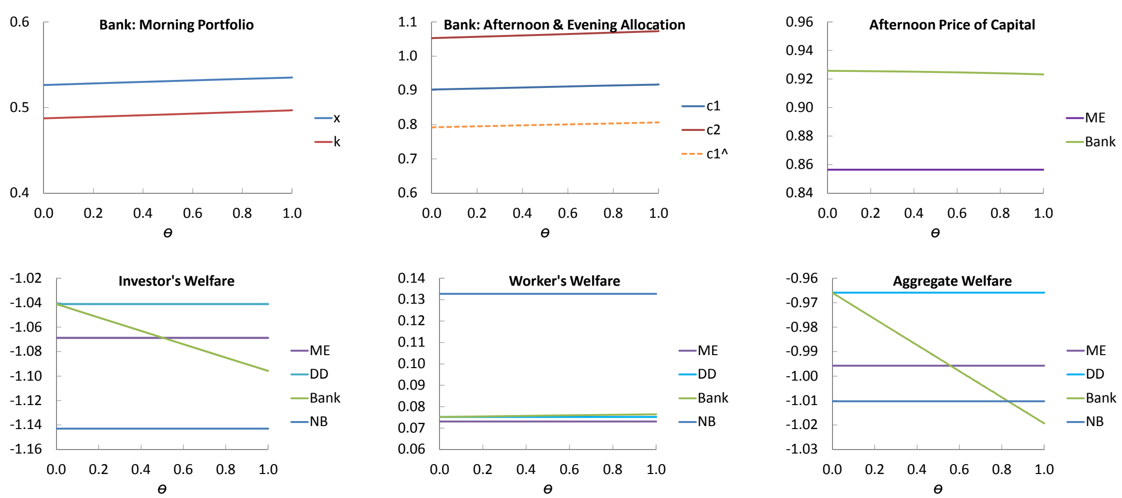


ME: monetary equilibrium; DD: no-run equilibrium; Bank: sunspot fractional banking; NB: narrow banking.

prefer whichever system creates a higher demand for cash: narrow banking is always best, while fractional reserve banking subject to runs dominates the no-run equilibrium and the monetary equilibrium.

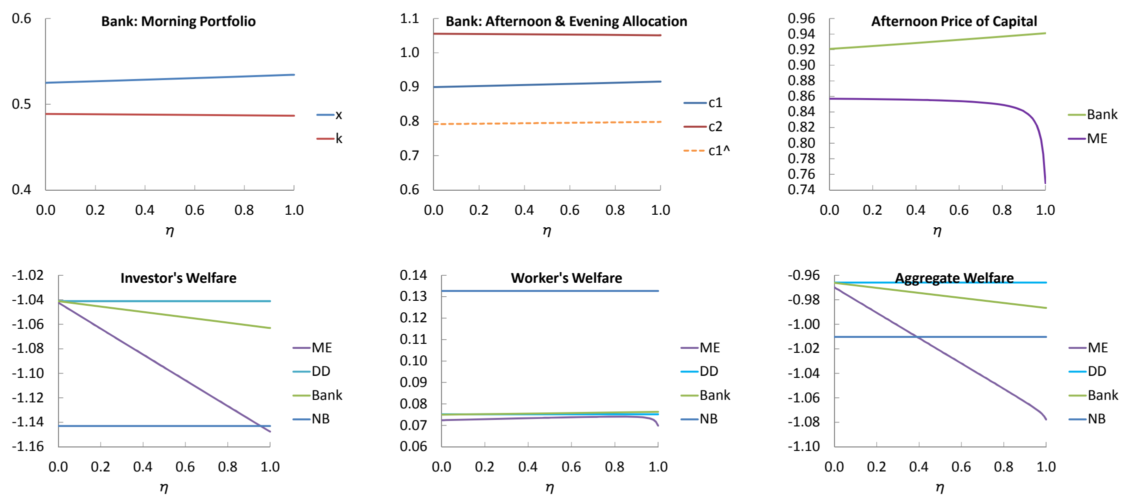
Finally, Figure 3 shows the effect of increasing the frequency of market freezes. In this case, in a fractional reserve banking system, money and afternoon allocations are increasing in η, while capital and evening consumption are decreasing in η. Basically, more frequent

Figure 2: Frequency of coordination failure



ME: monetary equilibrium; DD: no-run equilibrium; Bank: sunspot fractional banking; NB: narrow banking.

Figure 3: Frequency of market freeze



ME: monetary equilibrium; DD: no-run equilibrium; Bank: sunspot fractional banking; NB: narrow banking.

market freezes makes capital less useful. Note, however, that conditional on the secondary market being open, the value of capital increases in η; the reason being there is more more money to be exchanged for less capital. Investor's welfare decreases with η since banks value goes down with more frequent market freezes.

7 Investment Risk [preliminary]

Suppose now that he rate of return on capital that matures in the evening is subject to risk. For simplicity, assume that the realized return falls in a two-point set, $R_t \in \{0, R\}$ with $\delta \equiv \Pr[R_t = 0] > 0$ and $(1-\delta)R > 1$. Assume that R_t is not known until it is realized in the evening.[15] Below, we provide a brief overview of the results we obtain in this case.

7.1 Optimal allocation

The planner's problem is now

$$\max\left\{-x - k + \pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right]\right\} \tag{50}$$

subject to

$$\begin{aligned} x - \pi c_1 &\geq 0 & (51)\\ R_t k + [x - \pi c_1] - (1-\pi)c_2(R_t) &\geq 0 \text{ for } R_t \in \{0, R\} & (52)\\ \delta u(c_2(0)) + (1-\delta)u(c_2(R)) - u(c_1) &\geq 0 & (53) \end{aligned}$$

[15] Our information structure differs from Allen and Gale (1998), who assume that R_t is revealed in the afternoon.

The conditions that characterize the optimum are given by

$$\begin{aligned} u'\left(c_1^*\right) &= 1 \\ \delta u'\left(c_2^*(0)\right) + (1-\delta)u'\left(c_2^*(R)\right) &= 1 \\ (1-\delta)Ru'\left(c_2^*(R)\right) &= 1 \end{aligned} \tag{54}$$

with $x^* = \pi c_1^* + (1-\pi)c_2^*(0)$ and $k^* = (1-\pi)\left[c_2^*(R) - c_2^*(0)\right]/R$. Note that an efficient allocation satisfies, $x, k > 0$ and $c_2^*(0) < c_1^* < c_2^*(R)$. The following result follows.

Lemma 9 *If preferences exhibit decreasing absolute risk aversion, the incentive constraint (53) does not bind in the choice problem (50)-(53).*

Proof of Lemma 9. Denote λ_1 $[\lambda_2(0), \lambda_2(R)$ and $\lambda_3]$ the Lagrange multipliers for the inequalities (51) [(52) and (53)]. The first-order conditions are

$$\begin{aligned} x &: \quad -1 + (\lambda_1 + \lambda_2(0) + \lambda_2(R)) = 0 \\ k &: \quad -1 + R\lambda_2(R) = 0 \\ c_1 &: \quad \pi u'(c_1) - \pi(\lambda_1 + \lambda_2(0) + \lambda_2(R)) - \lambda_3 u'(c_1) = 0 \\ c_2(0) &: \quad (1-\pi)\,\delta u'(c_2(0)) - (1-\pi)\,\lambda_2(0) + \delta u'(c_2(0))\lambda_3 = 0 \\ c_2(R) &: \quad (1-\pi)\,(1-\delta)\,u'(c_2(R)) - (1-\pi)\,\lambda_2(R) + (1-\delta)\,u'(c_2(R))\lambda_3 = 0 \end{aligned}$$

Setting $\lambda_1 = \lambda_3 = 0$ and rearranging yields (54).

We now show that the incentive constraint (53) is nonbinding. Let $\Phi(R) \equiv \delta u(c_2^*(0)) + (1-\delta)u\left(c_2^*(R)\right)$. The incentive constraint is nonbinding if

$$\Phi(R) \geq u(c_1^*).$$

From (54), $c_2^*(R)$ is increasing, $c_2^*(R)$ is decreasing, and c_1^* is independent of R. Furthermore, note that when $(1-\delta)R = 1$, then, from (54), $c_1^* = c_2^*(0) = c_2^*(R)$. Hence, $\Phi(1/(1-\delta)) = u\left(c_1^*\right)$. Thus, if $\Phi(R)$ is increasing in R, the incentive constraint is strictly nonbinding for any $R > 1$. We now derive conditions under which $\Phi(R)$ is increasing in R.

The derivative of $\Phi(R)$ is

$$\Phi'(R) = \delta u'(c_2^*(0))\frac{dc_2^*(0)}{dR} + (1-\delta)u'\left(c_2^*(R)\right)\frac{dc_2^*(R)}{dR}. \tag{55}$$

From (54), we have

$$\frac{dc_2^*(R)}{dR} = -\frac{1}{(1-\delta)R^2 u''\left(c_2^*(R)\right)} \text{ and } \frac{dc_2^*(0)}{dR} = \frac{1}{\delta R^2 u''\left(c_2^*(0)\right)}.$$

Substituting these expressions into (55), yields

$$\Phi'(R) = \frac{1}{A\left(c_2^*(0)\right)R^2}\left[\frac{A\left(c_2^*(0)\right)}{A\left(c_2^*(R)\right)} - 1\right],$$

where $A(c) = -u''(c)/u'(c)$ is the coefficient of absolute risk aversion. With decreasing absolute risk aversion, we have $A(c_2^*(0)) > A(c_2^*(R))$, implying $\Phi'(R) > 0$. The intuition is that with decreasing absolute risk aversion an exogenous increase in R increases wealth, and so the planner invests more into the risky asset.[16] $\blacksquare$

7.2 Monetary equilibrium

Evening consumption is now constrained by

$$c_2(0) \leq \frac{m_t}{p_t^a} - \xi k^P \tag{56}$$

$$c_2(R) \leq \frac{m_t}{p_t^a} - \xi k^P + R\left(k + k^P\right) \tag{57}$$

Combine (10) and (11) set to equality, together with (9), to form $c_1 = (\beta/\mu)x + \xi k^I + \omega\left(k - k^I\right)$. Similarly, we can transform (56) and (57) into $c_2(0) = (\beta/\mu)x - \xi k^P$ and $c_2(R) = (\beta/\mu)x - \xi k^P + R\left(k + k^P\right)$, respectively.

If the market is not open and the investor turns out to be patient, he carries m_t/p_t^a units of afternoon goods into the evening so that consumption is constrained by

$$\hat{c}_2(0) \leq \frac{m_t}{p_t^a} \tag{58}$$

$$\hat{c}_2(R) \leq \frac{m_t}{p_t^a} + Rk \tag{59}$$

Combine (10) and (13) set to equality, together with (9), to form $\hat{c}_1 = (\beta/\mu)x + \xi k$. Similarly, we can transform (58) and (59) into $\hat{c}_2(0) = (\beta/\mu)x$ and $\hat{c}_2(R) = (\beta/\mu)x + Rk$, respectively.

The investor's objective function satisfies

$$\max \quad \begin{array}{l} -x - k + (1-\eta)\left\{\pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right]\right\} \\ +\eta\left\{\pi u(\hat{c}_1) + (1-\pi)\left[\delta u(\hat{c}_2(0)) + (1-\delta)u(\hat{c}_2(R))\right]\right\} \end{array} \tag{60}$$

$$\text{s.t. } k - k^I \geq 0$$

The constraint means that the impatient investor cannot sell more capital than he holds. Denote the Lagrange multiplier for this constraint λ_k. Note that there is a similar constraint on cash for the patient investor. However, since Inada conditions hold, he will never sell all his cash, and so we ignore it. Furthermore, investors always carry cash into the afternoon since it is the only way to insure against the event that they are patient and $R_t = 0$. When $(1-\eta)\left[\pi\xi + (1-\pi)(1-\delta)R\right] + \eta\omega > \beta/\mu$, the expected return on capital is higher than the return on money and thus, investors will also hold some capital.

[16] Decreasing absolute risk aversion is widly assumed in the profession and there is a lot of experimental and empirical evidence that is consistent with it (see e.g. Friend, Irwin and Blume, Marshall (1975). "The Demand for Risky Assets". American Economic Review 65 (5): 900–922.)

Let $\hat{u}' = \pi u'(\hat{c}_1) + (1-\pi)\left[\delta u'(\hat{c}_2(0)) + (1-\delta)u'(\hat{c}_2(R))\right]$ and $\check{u}' = \pi\omega u'(\hat{c}_1) + (1-\pi)(1-\delta)Ru'(\hat{c}_2(R))$. An investor's desired portfolio choice is characterized by

$$\begin{aligned}
(1-\eta)\left\{\pi u'(c_1) + (1-\pi)\left[\delta u'(c_2(0)) + (1-\delta)u'(c_2(R))\right]\right\} + \eta\hat{u}' &= \mu/\beta &(61)\\
(1-\eta)\left[\pi\omega u'(c_1) + (1-\pi)(1-\delta)Ru'(c_2(R))\right] + \eta\check{u}' + \lambda_k &= 1 &(62)\\
(1-\eta)\,\pi u'(c_1)(\xi-\omega) &= \lambda_k &(63)\\
\delta u'(c_2(0))\,\xi + (1-\delta)u'(c_2(R))(\xi - R) &= 0 &(64)
\end{aligned}$$

We have the endogenous variables $(x, k, k^I, k^P, \lambda_k, \xi)$. Equations (61)-(64), the Kuhn-Tucker condition $\lambda_k(k-k^I) = 0$, and the market clearing condition

$$\pi k^I = (1-\pi)\,k^P$$

solve for these. Note that from (63), in any monetary equilibrium $\xi \geq \omega$ (if the market is open).

Substituting λ_k, yields the following first-order conditions:

$$\begin{aligned}
(1-\eta)\left\{\pi u'(c_1) + (1-\pi)\left[\delta u'(c_2(0)) + (1-\delta)u'(c_2(R))\right]\right\} + \eta\hat{u}' &= \mu/\beta &(65)\\
(1-\eta)\left[\pi u'(c_1)\,\xi + (1-\pi)(1-\delta)Ru'(c_2(R))\right] + \eta\check{u}' &= 1 &(66)\\
\delta u'(c_2(0))\,\xi + (1-\delta)u'(c_2(R))(\xi - R) &= 0 &(67)
\end{aligned}$$

Lemma 10 *Assume, $\omega \leq \beta/\mu$, then in any monetary equilibrium $(1-\delta)R \geq \xi \geq \omega$, where*

$$\xi = \frac{1-\eta\check{u}'}{\mu/\beta - \eta\hat{u}'}$$

with $\xi = \beta/\mu$ when $\eta = 0$.

Proof of Lemma 10. Rewrite (67) as follows:

$$\delta u'(c_2(0))\,\xi + (1-\delta)u'(c_2(R))\,\xi = (1-\delta)u'(c_2(R))\,R$$

Use (66) to substitute $(1-\delta)u'(c_2(R))\,R$ to get

$$\delta u'(c_2(0)) + (1-\delta)u'(c_2(R)) = \frac{1-(1-\eta)\,\pi u'(c_1)\,\xi - \eta\check{u}'}{\xi(1-\pi)(1-\eta)}$$

Replace $\delta u'(c_2(0)) + (1-\delta)u'(c_2(R))$ in (65) to get

$$\frac{1-\eta\check{u}'}{\xi} + \eta\hat{u}' = \mu/\beta$$

Simplify to get

$$\xi = \frac{1-\eta\check{u}'}{\mu/\beta - \eta\hat{u}'}$$

which establishes the result $\xi = \beta/\mu$ when $\eta = 0$. From (67), we get

$$\xi = \frac{(1-\delta)Ru'\left(c_2(R)\right)}{\delta u'\left(c_2(0)\right) + (1-\delta)u'\left(c_2(R)\right)}$$

Then, the inequality $u'\left(c_2(0)\right) > u'\left(c_2(R)\right)$, yields

$$\xi \leq (1-\delta)R.$$

$\xi \geq \beta/\mu$ since otherwise nobody would hold money. ■

Note that Lemma 10 and market clearing imply that for $\eta = 0$, $c_1 = (\beta/\mu)(x+k)$, $c_2(0) = (\beta/\mu)\left(x - \pi(1-\pi)^{-1}k\right)$, and $c_2(R) = (\beta/\mu)\left(x - \pi(1-\pi)^{-1}k\right) + Rk(1-\pi)^{-1}$. Furthermore, Lemma 10 and the market clearing condition imply

Proposition 11 *For $\eta = 0$, a monetary equilibrium is a (x, k) that satisfies*

$$\begin{aligned}
\pi u'(c_1) + (1-\pi)\left[\delta u'\left(c_2(0)\right) + (1-\delta)u'\left(c_2(R)\right)\right] &= \mu/\beta \\
\pi u'(c_1) + (1-\pi)(1-\delta)Ru'\left(c_2(R)\right)\mu/\beta &= \mu/\beta
\end{aligned}$$

where $c_1 = (\beta/\mu)(x+k)$, $c_2(0) = (\beta/\mu)\left(x - \pi(1-\pi)^{-1}k\right)$, and $c_2(R) = c_2(0) + Rk(1-\pi)^{-1}$.

Proof. There are two cases. Either the constraint binds or it does not bind. If it does not bind, then $\lambda_k = 0$ and so, from (63), $\xi = \omega$. Then, from Lemma 10, $\beta/\mu = \omega$. Accordingly, the constraint does not bind if and only if $\beta/\mu = \omega$ which is a degenerate case. Note that for $\beta/\mu < \omega$ no monetary equilibrium exists since no agents would acquire money.

If it binds, then $\lambda_k > 0$ and $k = k^I$. Furthermore, from (63), $\xi = \beta/\mu > \omega$. The equilibrium equations are derived by replacing $\xi = \beta/\mu$ in (65) and (66) and the consumption quantities. ■

The initial portfolio (x, k) constitutes part of a stationary monetary equilibrium. To recover equilibrium nominal values, impose the equilibrium condition $m_t = M_t$. From condition (10) we then have an expression for the morning price-level, $p_t^m = M_t/x$. Condition (9) then implies $p_t^a = (\mu/\beta)\,p_t^m$, and so on.

7.3 Diamond-Dybvig bank

Consider again a fractional reserve banking system not subject to runs. In this case, a bank holds cash reserves in sufficient quantities to meet expected redemptions. Any cash left over following afternoon redemption activity $[(\beta/\mu)\,x - \pi c_1]$ is used to purchase afternoon output and then stored to finance consumption in the evening. The bank therefore faces an evening state-contingent budget constraint

$$R_t k + [(\beta/\mu)\,x - \pi c_1] \geq (1-\pi)c_2(R_t) \text{ for } R_t \in \{0, R\}. \tag{68}$$

The bank's choice problem is to maximize

$$\max\left\{-x - k + \pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right]\right\}. \tag{69}$$

subject to (25), (68) and the incentive constraint (53).

The solution to this problem entails holding "excess reserves," that is, constraint (25) remains slack. As already mentioned, excess cash is used to purchase output in the afternoon to be stored for the evening. The purpose of this is to provide insurance in the evening against a poorly performing investment. The constraints (68) on the other hand, can be expected to bind tightly. Finally, in the proof of Lemma 12, we show that the incentive constraint (53) is nonbinding.

For a given monetary policy μ, the equilibrium consumption allocation is characterized by

$$\begin{aligned} u'(c_1) &= \mu/\beta \\ \delta u'(c_2(0)) + (1-\delta)\, u'(c_2(R)) &= \mu/\beta \\ (1-\delta)\, R u'(c_2(R)) &= 1 \end{aligned} \tag{70}$$

with

$$x = (\mu/\beta)\,[\pi c_1 + (1-\pi) c_2(0)] \tag{71}$$

$$k = (1-\pi)\,[c_2(R) - c_2(0)]\,/R \tag{72}$$

Lemma 12 *If preferences exhibit decreasing absolute risk aversion, the incentive constraint (53) is nonbinding in the bank's choice problem (69).*

Proof of Lemma 12. Denote λ_1 $[\lambda_2(0), \lambda_2(R)$ and $\lambda_3]$ the Lagrange multipliers for the inequalities (25) [(68) and (53)]. The first-order conditions are

$$\begin{aligned} x &: \; -1 + (\beta/\mu)\,(\lambda_1 + \lambda_2(0) + \lambda_2(R)) = 0 \\ k &: \; -1 + R\lambda_2(R) = 0 \\ c_1 &: \; \pi u'(c_1) - \pi(\lambda_1 + \lambda_2(0) + \lambda_2(R)) - \lambda_3 u'(c_1) = 0 \\ c_2(0) &: \; (1-\pi)\,\delta u'(c_2(0)) - (1-\pi)\,\lambda_2(0) + \delta u'(c_2(0))\lambda_3 = 0 \\ c_2(R) &: \; (1-\pi)\,(1-\delta)\,u'(c_2(R)) - (1-\pi)\,\lambda_2(R) + (1-\delta)\,u'(c_2(R))\lambda_3 = 0 \end{aligned}$$

Setting $\lambda_1 = \lambda_3 = 0$ and rearranging yields (70).

We now show that the incentive constraint (53) is nonbinding. Let $\tilde{\Phi}(\mu/\beta) \equiv \delta u(c_2(0)) + (1-\delta) u(c_2(R)) - u(c_1)$. The incentive constraint is nonbinding if

$$\Phi(\mu/\beta) \geq 0$$

From (70), $c_2(R)$ is independent and $c_2(R)$ and c_1 are decreasing in μ/β. Furthermore, note that when $\mu/\beta = 1$, Lemma (9) applies. Thus, if $\tilde{\Phi}(\mu/\beta)$ is increasing in μ/β, the incentive constraint is strictly nonbinding for any $\mu/\beta > 1$. We now derive conditions under which $\tilde{\Phi}(\mu/\beta)$ is increasing in μ/β.

The derivatives of $\tilde{\Phi}(\mu/\beta)$ is

$$\tilde{\Phi}'(\mu/\beta) = \delta u'(c_2(0)) \frac{dc_2(0)}{d(\mu/\beta)} - u'(c_1) \frac{dc_1}{d(\mu/\beta)}. \tag{73}$$

From (70), we have

$$\frac{dc_1}{d\left(\mu/\beta\right)} = \frac{1}{u''\left(c_1\right)} \text{ and } \frac{dc_2(0)}{d\left(\mu/\beta\right)} = \frac{R(R-1)^{-1}}{u''\left(c_2(0)\right)}.$$

Substituting these expressions into (55), yields

$$\tilde{\Phi}'\left(\mu/\beta\right) = \frac{1}{A\left(c_2(0)\right)}\left[\frac{A\left(c_2(0)\right)}{A\left(c_1\right)} - \frac{\delta R}{R-1}\right].$$

where $A\left(c\right) = -u''\left(c\right)/u'\left(c\right)$ is the coefficient of absolute risk aversion. With decreasing absolute risk aversion $A\left(c_2(0)\right) > A\left(c_1\right)$. Hence, since $\frac{\delta R}{R-1} < 1$, $\tilde{\Phi}'\left(\mu/\beta\right) > 0$. ■

Comparing (70) to (7), we see that the allocations correspond under the Friedman rule, $\mu = \beta$.

Lemma 13 *$x > 0$, $k > 0$ and $c_2(0) < c_1 < c_2(R)$.*

Proof. From (71), $x = 0$ implies $c_1 = c_2(0) = 0$; $x > 0$ follows from $u'(0) = \infty$. From (72), $k = 0$ implies $c_2(0) = c_2(R) = c_2$. From (70), we get $u'(c_2) = \mu/\beta$ and $(1-\delta)Ru'(c_2) = 1$. Thus, $\beta/\mu = (1-\delta)R$, which is a contradiction since $\beta/\mu \leq 1$ and $(1-\delta)R > 1$.

Given $\mu/\beta \geq 1$ and $(1-\delta)R > 1$, (70) implies $u'(c_1) = \mu/\beta > [(1-\delta)R]^{-1} = u'(c_2(R))$. Thus, $c_1 < c_2(R)$. Again from (70) we get $u'(c_1) = \delta u'(c_2(0)) + (1-\delta)u'(c_2(R))$. If $c_1 = c_2(0)$, then $c_1 = c_2(R)$, a contradiction. If $c_1 < c_2(0)$ then $u'(c_1) < u'(c_2(R))$ and so $c_1 > c_2(R)$, a contradiction. Thus, $c_1 > c_2(0)$. Finally, (72) implies $k > 0$. ■

7.4 Fractional reserve banks subject to runs and market freezes

We now revisit a fractional reserve banking system subject to runs. Given risky investment, selling capital in the secondary market when there is a run implies

$$\tilde{c}_2(0) \leq (\beta/\mu)x + \xi k - \xi k^P \tag{74}$$
$$\tilde{c}_2(R) \leq (\beta/\mu)x + \xi k - \xi k^P + Rk^P \tag{75}$$

In the event of a run, we assume parameters such that full liquidation/selling of the bank's assets is insufficient to meet its short-term obligations, so that $\tilde{c}_1, \hat{c}_1 < c_1$ in any equilibrium (where $\tilde{c}_1$ is given by (31) and $\hat{c}_1$ by (33)). Also, because $\mu > \beta$, we anticipate that the bank will not want to carry cash into the following period. In this case, the bank's choice problem can be stated as follows

$$W^B(\theta,\mu) \equiv \max\left\{\begin{array}{l} -x - k + (1-\theta)\left\{\pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right]\right\} \\ +(1-\eta)\,\theta u\left[\pi u(\tilde{c}_1) + (1-\pi)\left[\delta u(\tilde{c}_2(0)) + (1-\delta)u(\tilde{c}_2(R))\right]\right] + \eta\theta u\left(\hat{c}_1\right) \end{array}\right\} \tag{76}$$

subject to (31), (33), (74), (75), the incentive constraint (53), and

$$\left(\beta/\mu\right)x \geq \pi c_1 \tag{77}$$
$$R_t k + \left[\left(\beta/\mu\right)x - \pi c_1\right] \geq (1-\pi)c_2(R_t) \text{ for } R_t \in \{0, R\} \tag{78}$$

Because $u'(0) = \infty$ and $R_t = 0$ with positive probability, we can safely assume that the cash reserve constraint (77) remains slack. Furthermore, we solve the problem without taking (53) into account and later verify numerically that it holds. Let $\tilde{u}' = \pi u'(\tilde{c}_1) + (1-\pi)\left[\delta u'(\tilde{c}_2(0)) + (1-\delta)u'(\tilde{c}_2(R))\right]$. The solution to the bank's problem is characterized by the following set of conditions

$$\begin{aligned}(1-\theta)\,u'(c_1) + \theta\left[(1-\eta)\,\tilde{u}' + \eta u'\left(\hat{c}_1\right)\right] - \mu/\beta = 0\\ \delta u'(c_2(0)) + (1-\delta)\,u'(c_2(R)) - u'(c_1) = 0\\ (1-\theta)\,(1-\delta)Ru'(c_2(R)) + \theta\left[(1-\eta)\,\xi\tilde{u}' + \eta\omega u'\left(\hat{c}_1\right)\right] - 1 = 0\\ \delta u'(\tilde{c}_2(0))\xi + (1-\delta)u'(\tilde{c}_2(R))\,(\xi - R) = 0\end{aligned} \tag{79}$$

where $c_2(0) = [(\beta/\mu)x - \pi c_1]/(1-\pi)$, $c_2(R) = [(\beta/\mu)x - \pi c_1 + Rk]/(1-\pi)$, $\tilde{c}_1 = (\beta/\mu)x + \xi k$, $\tilde{c}_2(0) = (\beta/\mu)x - \xi\pi\,(1-\pi)^{-1}\,k$, $\tilde{c}_2(R) = (\beta/\mu)x + (R - \xi\pi)\,k\,(1-\pi)^{-1}$ and $\hat{c}_1 = (\beta/\mu)x + \omega k$. These four first-order conditions plus market clearing solve for $\left(x, k, \xi, k^P, c_1\right)$.

7.5 Lender of last resort facility

In this section, we consider the costs and benefits of operating a lender-of-last-resort (LOLR) facility. In our model, early liquidation is triggered by a sunspot event. From an *ex ante* perspective, early liquidation is inefficient. One way to prevent it is to impose a 100% reserve requirement on all short-term obligations. But another way to prevent inefficient liquidation is to grant fractional reserve banks permission to access a lending facility operated by a central bank with the power to create cash–the object of redemption in privately-created demand deposit liabilities.

We first present a LOLR intervention that restores efficiency. What is remarkable about it is that it does not require commitment as, for example, a suspension clause.

7.6 Efficient LOLR intervention without commitment

To restore efficiency we need to assume that the LOLR provides a facility at the beginning of every afternoon. At this facility, banks are able to secure collaterlized loans from the LOLR.[17] Of course, in the model, this collateral is risky: it's expected return is $(1-\delta)Rk$. In many jurisdictions, central banks are prohibited from accepting high risk collateral in their emergency lending facility, but clearly, some risk is always present.

Efficiency requires that banks offer a specific state contingent contract in the afternoon.[18] Before we present this contract, let us be clear about the information flow and sequence of actions during the afternoon. First, the sunspot is observed. If the sunspot is on, all investors arrive early to the bank. The central bank then makes a collateralized loan to the private bank, where the loan size is sufficient to finance the short-term obligations c_1 to all

[17] In general, we could include a discount rate, but we abstract from discounting for now.

[18] In our set-up with a random capital return, the offer has to be state contingent, but this feature of the offer is not always needed. If the capital return is deterministic, the efficient offer is not state contigent. This shows that state contingency of the optimal offer is not essential for restoring efficiency.

investors. Since the bank holds $x\left(\beta/\mu\right) = (1-\pi)c_2(0) + \pi c_1$ units of real cash balances and has promised c_1 consumption to all agents, it requires a real loan of

$$\left[c_1 - x\left(\beta/\mu\right)\right] = (1-\pi)\left[c_1 - c_2(0)\right] \tag{80}$$

After receiving the loan, the bank pays out all cash to investors, fulfilling its short-term obligations without liquidating capital. Next, the bank makes the following state-contingent offer: in exchange for $p_t^a c_1$ units of cash it offers evening consumption $c_2(0)$ when $R_t = 0$ and $c_2(R)$ when $R_t = R$.[19] Note this offer is incentive-feasible since

$$u\left(c_1\right) \leq \delta u\left(c_2\left(0\right)\right) + \left(1-\delta\right) u\left(c_2\left(R\right)\right).$$

That is, all patient agents accept the offer and all impatient agents reject it. Note that this offer is optimal since it replicates the constrained efficient allocation. Furthermore, the offer has no effect on the stock of money in the economy. To see this, note that after the run, the total real quantity of money in the hand of investors is c_1. The bank then receives from the patient investors $(1-\pi)\,c_1$ units of real cash. Finally, it acquires $(1-\pi)c_2(0)$ afternoon goods and so the workers obtain $x\left(\beta/\mu\right) = (1-\pi)c_2(0) + \pi c_1$ units of money, which is the same quantity as without a run. Finally, the bank holds $(1-\pi)\left(c_1 - c_0\right)$ units of real cash. It then uses the cash to repay the loan at the end of the afternoon.[20]

It is important to note that efficiency can only be restored with the help of the LOLR. In the absence of a LOLR, the same state contingent offer does not work. It only works with a LOLR because the patient investors know that the bank is under no obligation of liquidating capital. This implies that patient investors have no risk that there is no payout when $R_t = R$.

Political economy In practice, perhaps for political economy reasons, the LOLR might not be willing to provide a loan with certainty. To capture this, we assume that the facility opens at the beginning of every afternoon period with probability $0 \leq \alpha \leq 1$. The idea here is that there may be some uncertainty as to whether the central bank is willing to extend emergency funding.[21]

If the facility is open, banks are able to borrow up to the full market value of their collateral as above. Under these assumptions, the bank's program is as follows:

$$\begin{aligned} &-x - k + \left(1-\tilde{\theta}\right)\left\{\pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right]\right\} \\ &+\tilde{\theta} u\left[(\beta/\mu)x + \rho k\right] \end{aligned} \tag{81}$$

[19] One could assume that the emergency loan is paid back only when economic fundamentals turn out to be good ($R_t = R$). When the return to capital is bad ($R_t = 0$), banks make no effort to repay their money loan and the central bank keeps the collateral. In effect, there is a helicopter drop of money to investors when fundamentals turn out to be bad. Such an arrangement, however, would not restore efficiency because of the distortionary effect of inflation.

[20] The loan is required to be pais back at the end of the afternoon to avoid a situation where $R_t = 0$ and the bank wants to keep the money.

[21] Consider, for example, the absence of Fed intervention in the Lehman Brothers' failure in November 2008 following its intervention in Bear Stearns in March 2008.

where $\tilde{\theta} = \theta(1-\alpha)$. It is clear from this that a change in α is equilvalent to a change in the sunspot probability θ. That is, there can be two reasons for a high $\tilde{\theta}$: The society can be inherent instable as reflected by a large θ or because there are less frequent LOLR interventions. This finding also yields an alternative interpretation for our notion of inherent fragility that we have discussed when we introduced the parameter θ. The alternative interpretation is that societies are equally prone to runs as captured by a common sunspot probability θ, but they differ in their political economy which is reflected in differences in α.

7.7 Making Creditors Whole: Bailouts

In what follows we consider a bailout policy. As we will see this policy is clearly suboptimal from a societies point of view but it is promoted by the investors (the banked agents in our model). To make it less obvious that it is a hidden subsidy for the investors we are going to design it in a way that it also prevents bank-runs. We proceed in two steps. We first consider a bailout policy that does not prevent bank-runs after which we design it such that it prevents also bank-runs.

Outright Bailouts The bailout policy consists of a cash transfer to patient investors. It is paid when the return on capital is low. Assume for a moment, that the cash transfer takes only place when there is no run. This modifies the expected payoff of investors as follows:

$$\begin{aligned}&-x-k+(1-\theta)\left\{\pi u(c_1)+(1-\pi)\left[\delta\left[u(c_2(0))+\beta b\right]+(1-\delta)u(c_2(R))\right]\right\} \\ &+\theta u\left[(\beta/\mu)x+\rho k\right],\end{aligned} \tag{82}$$

where b the utility from the nominal cash transfer $p^m_{t+1}b$. The money can only be spent the next morning which explains the discount factor.

It is clear that this cash transfer can not be too high. Otherwise, impatient investors misrepresent their type. This requires that

$$u(c_1) \geq \delta\beta b + u(0),$$

where here $u(0)\,(=0)$ is the utility of impatient agents that consume evening goods. Since these bailouts consist of lump-sum cash transfers, their allocative effects depends on whether the injected cash is neutralized (taken out) the following morning. If it is neutralized, for example by reducing the lump-sum transfers of cash to workers, then inflation does not change and the allocation does not change. The workers are clearly worse off with these bailouts.

If the cash transfers are not neutralized, then expected inflation increases which does effect all endogenous variables. This will have similar effects than increasing the deterministic inflation rate of the previous sections. There will be a reduction of the equilibrium x and an increase in the equilibrium k.

Bailout that prevent bank runs In what follows we study under which condition the cash subsidy in the evening can prevent a bank run as well. To prevent a run, it has to be designed such that a patient investors is willing to wait in the afternoon even if he believes that all other patient investors run. The following inequality captures this condition:

$$u\left[(\beta/\mu)x + \rho k\right] \leq u(0) + \beta b$$

The left-hand side of this inequality is the utility of a patient investor gets when he runs and all other patient investors run as well. The first term on the right-hand side is the utility from consumption in the evening. Since he expects that the bank liquidates the capital and so he expects that no consumption goods are available and so his utility is $u(0)$. The second term is the discounted utility from the nominal cash transfer $p_{t+1}^{m} b$.

We can derive a range for b such that no investor mirepresents himself. Assuming $u(0) = 0$, the range satisfies

$$\delta u\left[(\beta/\mu)x + \rho k\right] \leq \delta\beta b \leq u(c_1)$$

This condition is easily satisfied for small δ.

8 References

1. Allen, Franklin and Douglas Gale (1998). “Optimal Financial Crises,” *Journal of Finance* 53, 1245-1283.

2. Allen, Franklin and Douglas Gale (2004). Financial Intermediaries and Markets. *Econometrica* 72, 1023-1061.

3. Andolfatto, David (2013). “Incentive-Feasible Deflation,” *Journal of Monetary Economics*, 60(4): 383-390.

4. Andolfatto, David, Ed Nosal and Bruno Sultanum (2014). “Preventing Bank Runs,” Federal Reserve Bank of St. Louis working paper.

5. Arifovic, Jasmina, Jiang, Janet and Yiping Xu (2013). “Experimental Evidence of Bank Runs as Pure Coordination Failures,” *Journal of Economic Dynamics and Control* 37(12):2446-2465.

6. Bryant, John (1980). “A Model of Reserves, Bank Runs, and Deposit Insurance,” *Journal of Banking and Finance*, 4: 335-344.

7. Camous, Antoine and Russell Cooper (2014). “Monetary Policy and Debt Fragility,” NBER Working Paper No. 20650.

8. Champ, Bruce, Bruce D. Smith and Stephen D. Williamson (1996), “Currency Elasticity and Banking Panics: Theory and Evidence,” *Canadian Journal of Economics*, 29(4):828-864.

9. Cooper, Russell and Ross, T.W. (1991). “Bank runs: liquidity and incentives”, NBER Working Paper #3921, November.

10. Cooper, Russell and Ross, T.W. (1998). “Bank runs: liquidity costs and investment distortions”, *Journal of Monetary Economics* 41, 27–38.

11. Cooper, Russell and Kalin Nikolov (2013). “Government Debt and Banking Fragility: The Spreading of Strategic Uncertainty,” NBER Working Paper No. 19278.

12. Diamond, Douglas and Philip Dybvig (1983). “Bank Runs, Deposit Insurance, and Liquidity”, *Journal of Political Economy* 91(3): 401-419.

13. Ennis, Huberto M. and Todd Keister (2006). “Bank Runs and Investment Decisions Revisited”, *Journal of Monetary Economics* 53: 217-232.

14. Ennis, Huberto M. and Todd Keister (2003). “Economic growth, liquidity, and bank runs”, *Journal of Economic Theory* 109, 220–245.

15. Ennis, Huberto M. and Todd Keister (2009). ”Bank runs and institutions: The perils of intervention”, *American Economic Review* 99, 1588–1607.

16. Ennis, Huberto M. and Todd Keister (2010). ”Banking panics and policy responses”, *Journal of Monetary Economics* 57, 404–419.

17. Freixas, Xavier and Jean-Charles Rochet (1997), Microeconomics of Banking, MIT Press, Cambridge, MA.

18. Friend, Irwin and Marshall E. Blume (1975). "The Demand for Risky Assets", *American Economic Review* 65(5): 900-922.

19. Goldstein, Itay and Ady Pauzner (2005). "Demand Deposit Contracts and the Probability of Bank Runs, *Journal of Finance*, 60(3): 1293-1328.

20. Gorton, Gary (1988). "Banking Panics and Business Cycles," Oxford Economic Papers, 40: 751-781.

21. Jiang, Janet (2008). "Banking Crises in Monetary Economies," *Canadian Journal of Economics*, 41: 80-104.

22. Keister, Todd and Vijay Narasiman (2015). "Expectations vs. fundamentals-driven bank runs: When should bailouts be permitted?" *Review of Economic Dynamics*, Available online 14 March 2015.

23. Lagos, Ricardo and Randall Wright (2005). "A Unified Framework for Monetary Theory and Policy Analysis" *Journal of Political Economy* 113(3): 463-484.

24. Lai, Alexandra (2003). "Modelling Financial Instability: A Survey of the Literature," Bank of Canada working paper 2002-12.

25. Loewy, B. Michael (1991). "The Macroeconomic Effects of Bank Runs: An Equilibrium Analysis," *Journal of Financial Intermediation* 1: 242-256.

26. Peck, James and Karl Shell (2003). Equilibrium bank runs. *Journal of Political Economy* 111, 103–123.

27. Rochet, Jean-Charles and Xavier Vives (2004). "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?" *Journal of the European Economic Association*, 2(6):1116-1147.

28. Smith, D. Bruce (2003). "Taking Intermediation Seriously," *Journal of Money, Credit, and Banking* 35:1319-1357.

29. Wallace, Neil (1990). A banking model in which partial suspension is best. Federal Reserve Bank of Minneapolis Quarterly Review 14 (Fall), 11–23.