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

September 2015

Abstract

This paper integrates the Diamond and Dybvig (1983) theory of financial fragility with the Lagos and Wright (2005) model of monetary exchange. Our theory suggests that financial fragility is more likely to emerge in economies characterized by a low degree of social fragility. Where the propensity for coordination failure is high, narrow banking systems dominate fractional reserve banking systems. High inflation economies penalize narrow banking systems relatively more than fractional systems. In this way, high inflation promotes financial instability. Lendor-of-last resort policies financed by inflation can promote financial stability in fractional reserve systems, but our model suggests that not all segments of society benefit from the policy.

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.

[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.

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 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-

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

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.

It is interesting to report, however, that the workers in our model do not necessarily have their interests aligned with investors. For reasons that we still need to flesh 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 large extant literature built around the original Diamond and Dybvig (1983) paper; see Lai (2002) for a useful survey. Our paper is related to Allen and Gale (1998) as well as other authors who also model business cycle risk in capital returns. However, unlike that branch of the literature, there is no "leading indicator" or "news" that permits banks to forecast an increased likelihood of deteriorating fundamentals. Thus, in our model, bank runs are only triggered by psychological factors–as in the original Diamond and Dybvig (1983) model.[4]

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 in economic fundamentals.

We use our model to the investigate the desirability and consequences of lender-

[4]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.

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 liquidated in the afternoon is $0 < \xi < 1$.[5] The rate of return on capital liquidated 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.[6] Finally, assume that unliquidated capital depreciates fully at the end of the evening.[7]

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. Since R_t is known in the evening, the evening allocation can be conditioned on R_t. Since the problem is static, it may be written as

$$\max\left\{-x - k + \pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right]\right\} \tag{3}$$

subject to

$$\begin{aligned} y_1 &\geq \pi c_1 && (4) \\ R_t k + [y_1 - \pi c_1] &\geq (1-\pi)c_2(R_t) \text{ for } R_t \in \{0, R\} && (5) \\ x - y_1 &\geq v && (6) \end{aligned}$$

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).

We now characterize an efficient allocation. To begin, we can deduce that the constraint (4) will remain strictly slack. The reason for this is because it will be prudent to store some output from the afternoon to the evening in the event that the investment fails to pay off. The constraints (5) and (6), on the other hand, can be expected to bind tightly. The conditions that characterize the optimum are given by

$$\begin{aligned} u'\left(c_1^*\right) &= 1 \\ \delta u'\left(c_2^*(0)\right) + (1-\delta)u'\left(c_2^*(R)\right) &= 1 \\ (1-\delta)Ru'\left(c_2^*(R)\right) &= 1 \end{aligned} \tag{7}$$

With the consumption allocation so determined, use (5) for $R_t = 0$ to determine $y_1^* = \pi c_1^* + (1-\pi)c_2^*(0)$. Now use this latter condition and (5) for $R_t = R$ to determine $k^* = (1-\pi)\left[c_2^*(R) - c_2^*(0)\right]/R$. Finally, use (6) to determine $x^* = v + y_1^*$.

[5] Technically, the capital is "scrapped" and consumed instead of "liquidated" in the sense of converting to cash.

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

[7] It would be of some interest to permit durable capital with $R(t)$ representing a dividend, but for now we want to keep things as simple as possible.

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.[8]

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.[9] 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.[10] Thus, the quantity

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

[8]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.

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

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

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

For simplicity, assume that there is no secondary asset market in the afternoon that would permit patient and impatient investors to trade afternoon output for claims to evening output.[11] In this set up then, the capital investment associated with an impatient investor needs to be liquidated. If the investor turns out to be impatient, he faces the expenditure constraint

$$c_1 \leq \frac{m_t}{p_t^a} + \xi k \tag{11}$$

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

$$c_2(0) \quad \leq \quad \frac{m_t}{p_t^a} \tag{12}$$

$$c_2(R) \quad \leq \quad \frac{m_t}{p_t^a} + Rk \tag{13}$$

Combine (10) and (11) set to equality, together with (9), to form $c_1 = (\beta/\mu)x + \xi k$. Similarly, we can transform (12) and (13) into $c_2(0) = (\beta/\mu)x$ and $c_2(R) = (\beta/\mu)x + Rk$, respectively. Substituting these variables into the investor's objective function implies

$$\max \left\{ \begin{array}{l} -x - k + \pi u[(\beta/\mu)x + \xi k] \\ +(1-\pi)\left[\delta u\left((\beta/\mu)x\right) + (1-\delta)u\left((\beta/\mu)x + Rk\right)\right] \end{array} \right\} \tag{14}$$

Lemma 1 *$x > 0$; if $(\mu/\beta)[\pi\xi + (1-\pi)(1-\delta)R] - 1 > 0$ then $k > 0$.*

Proof. Since $c_2(0) = (\beta/\mu)x$, $x > 0$ follows from $u'(0) = \infty$. To show $k > 0$, take the first-order condition of problem (14) with respect to k including the non-negativity constraint $k \geq 0$ with Lagrange multiplier ζ_k. We obtain: $-1 + \pi\xi u'(c_1) + (1-\pi)(1-\delta)Ru'(c_2(R)) + \zeta_k = 0$. If $k = 0$ then $c_1 = c_2(0) = c_2(R) = c$ and thus, $\zeta_k = 1 - u'(c)[\pi\xi + (1-\pi)(1-\delta)R] \geq 0$. Furthermore, from the first-order condition with respect to x, we obtain $u'(c) = \mu/\beta$. Thus, $\zeta_k = 1 - (\mu/\beta)[\pi\xi + (1-\pi)(1-\delta)R] \geq 0$. Hence, if $(\mu/\beta)[\pi\xi + (1-\pi)(1-\delta)R] - 1 > 0$, $\zeta_k < 0$, a contradiction, and thus, $k > 0$. ■

Assuming parameters are such that $k > 0$, an investor's desired portfolio choice is characterized by

$$\begin{array}{l} \pi u'\left((\beta/\mu)x + \xi k\right) \\ +(1-\pi)\left[\delta u'\left((\beta/\mu)x\right) + (1-\delta)u'\left((\beta/\mu)x + Rk\right)\right] = \mu/\beta \end{array} \tag{15}$$

[11] Our purpose here is simply to demonstrate how money functions in this economy. Permitting a secondary market would complicate the exposition without altering the intuition. And later, when we study banking arrangements, the secondary market is not needed, so we abstract from it here.

$$\pi\xi u'\left((\beta/\mu)x+\xi k\right)+(1-\pi)(1-\delta)Ru'\left((\beta/\mu)x+Rk\right)=1 \tag{16}$$

As we showed in the proof of Lemma 1 when parameters are such that $k=0$, then the monetary equilibrium features $c_1=c_2(0)=c_2(R)=(\beta/\mu)x$ solving $u'((\beta/\mu)x)=\mu/\beta$.

The portfolio (x,k) that satisfies (15) and (16) 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 liquidates (consumes) his capital. A patient investor also spends all his money for afternoon goods, storing them into the evening. In the evening, the patient investor consumes his stored goods and the return to the maturing investment.

Comparing (15) and (16) to (7) reveals that the monetary equilibrium is not efficient under any monetary policy μ. This is not surprising as risk-averse investors are left to self-insure in the monetary economy and, moreover, capital is occasionally scrapped because of the assumed lack of a secondary market.

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've 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. 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

$$(\beta/\mu)\, x \geq \pi c_1 \tag{17}$$

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.[12] The bank therefore faces an evening state-contingent budget constraint

$$R_t k + [(\beta/\mu)\, x - \pi c_1] \geq (1-\pi) c_2(R_t) \text{ for } R_t \in \{0, R\} \tag{18}$$

The bank's choice problem is to maximize

$$\max\left\{-x - k + \pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta) u(c_2(R))\right]\right\} \tag{19}$$

subject to (17) and (18).

The solution to this problem entails holding "excess reserves," that is, constraint (17) 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. For a given monetary policy μ, the equilibrium consumption allocation is characterized by

$$\begin{aligned} u'(c_1) &= \mu/\beta \\ \delta u'(c_2(0)) + (1-\delta)\, u'(c_2(R)) &= \mu/\beta \\ (1-\delta)\, R u'(c_2(R)) &= 1 \end{aligned} \tag{20}$$

[12] 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.

with

$$x = (\mu/\beta)\left[\pi c_1 + (1-\pi)c_2(0)\right] \quad (21)$$
$$k = (1-\pi)\left[c_2(R) - c_2(0)\right]/R \quad (22)$$

Comparing (20) to (7), we see that the allocations correspond under the Friedman rule, $\mu = \beta$.

Lemma 2 *$x > 0$, $k > 0$ and $c_2(0) < c_1 < c_2(R)$.*

Proof. From (21), $x = 0$ implies $c_1 = c_2(0) = 0$; $x > 0$ follows from $u'(0) = \infty$. From (22), $k = 0$ implies $c_2(0) = c_2(R) = c_2$. From (20), 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 $(1-\delta)R > 1$, (20) implies $u'(c_1) = \mu/\beta > [(1-\delta)R]^{-1} = u'(c_2(R))$. Thus, $c_1 < c_2(R)$. Again from (20) 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, $c_1 > c_2(0)$. ■

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 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.[13] 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.[14] Note that the spirit of

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

[14]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.

sequential service is preserved here in the sense that patient investors will not be serviced if they fail to arrive early (in the afternoon).[15]

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 (17) 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 $c_1 \leq (\beta/\mu)x$. This latter condition, together with (21) implies

$$c_1 \leq (\beta/\mu)x = [\pi c_1 + (1-\pi)c_2(0)] \tag{23}$$

Lemma 2 implies that condition (23) is violated in the event of a run. In order to meet its short-term obligations then, the bank will have to liquidate capital. The maximum amount of purchasing power the bank can dispense through liquidation is given by

$$\tilde{c}_1 = (\beta/\mu)x + \xi k \tag{24}$$

Note that it is possible here that $\tilde{c}_1 > c_1$, in which case the bank need only liquidate some of its capital to honor its redemption promise. For the purpose of explaining the basic idea, we can assume here that liquidation is very costly (ξ sufficiently close to zero) so that $\tilde{c}_1 < c_1$ 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. In this case, it is immediately evident that a bank run equilibrium exists.

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.

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

[15] 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.

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 of the bank's assets is insufficient to meet its short-term obligations, so that $\tilde{c}_1 < c_1$ in any equilibrium (where $\tilde{c}_1$ is given by 24). 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

$$W^B(\theta,\mu) \equiv \max \left\{ \begin{array}{l} -x - k + (1-\theta)\left\{\pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right]\right\} \\ +\theta u\left[(\beta/\mu)x + \xi k\right] \end{array} \right\} \quad (25)$$

subject to

$$\begin{aligned} (\beta/\mu)\, x &\geq \pi c_1 && (26) \\ R_t k + \left[(\beta/\mu)\, x - \pi c_1\right] &\geq (1-\pi)c_2(R_t) \text{ for } R_t \in \{0, R\} && (27) \end{aligned}$$

Note that in the event of a run, depositors are better off *ex post* having their capital liquidated for some strictly positive return in the afternoon relative having their capital investment vanish when fundamentals turn out to be bad.

Because $u'(0) = \infty$ and $R_t = 0$ with positive probability, we can safely assume that the cash reserve constraint (26) remains slack. The solution to the bank's problem is characterized by the following set of conditions

$$\begin{array}{c} (1-\theta)\, u'(c_1) + \theta u'\left(\tilde{c}_1\right) = \mu/\beta \\ \delta u'(c_2(0)) + (1-\delta)\, u'(c_2(R)) = u'(c_1) \\ (1-\theta)\,(1-\delta)\, R u'(c_2(R)) + \theta\xi u'\left(\tilde{c}_1\right) = 1 \end{array} \quad (28)$$

where $c_2(0) = [(\beta/\mu)x - \pi c_1]/(1-\pi)$, $c_2(R) = [(\beta/\mu)x - \pi c_1 + Rk]/(1-\pi)$ and $\tilde{c}_1 = (\beta/\mu)x + \xi k$.

In comparing (28) to (20), we see that the possibility of bank runs θ and their associated liquidation costs ξ introduce "wedges" that alter the bank's optimal portfolio choice, relative to the case in which runs are absent ($\theta = 0$).

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:

$$\beta = 0.90 \quad \mu = 1.05 \quad \sigma = 1.50 \quad \delta = 0.10 \quad R = 1.20 \quad \theta = 0.10 \quad \xi = 0.60 \quad \pi = 0.50 \quad (29)$$

In what follows, we report results for how the allocation varies with the frequency of the sunspot (θ), the frequency of the productivity shock (δ), and monetary policy

Figure 1: Frequency of coordination failure

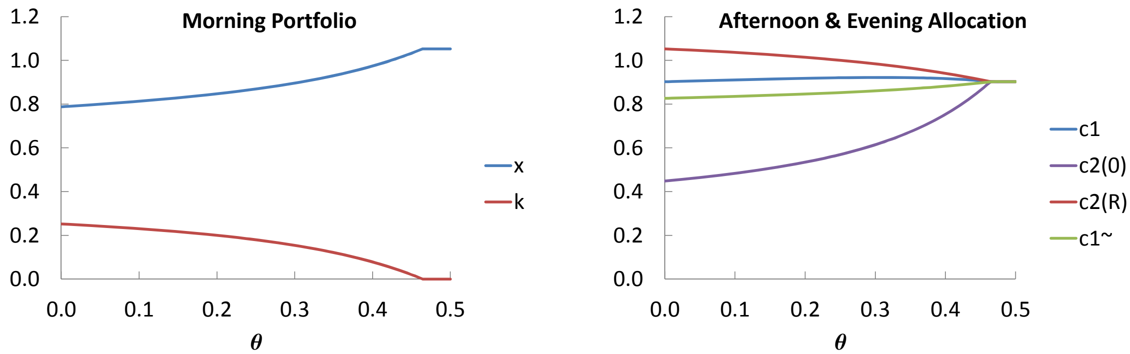


Figure 2: Frequency of bad fundamentals

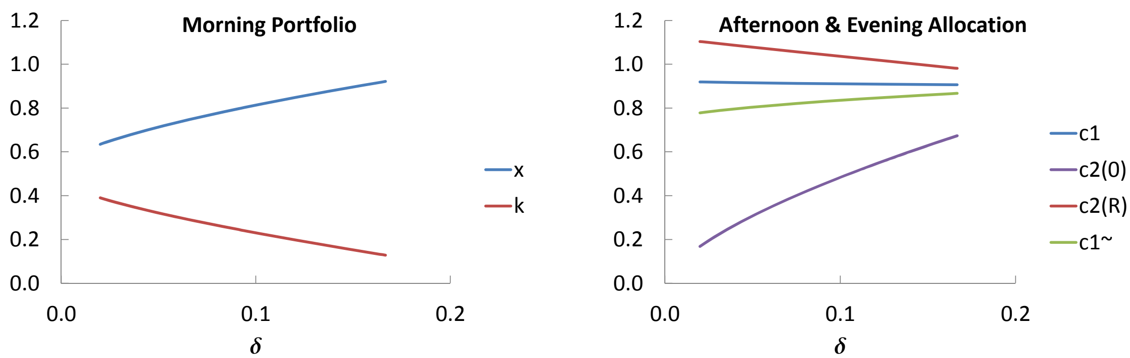
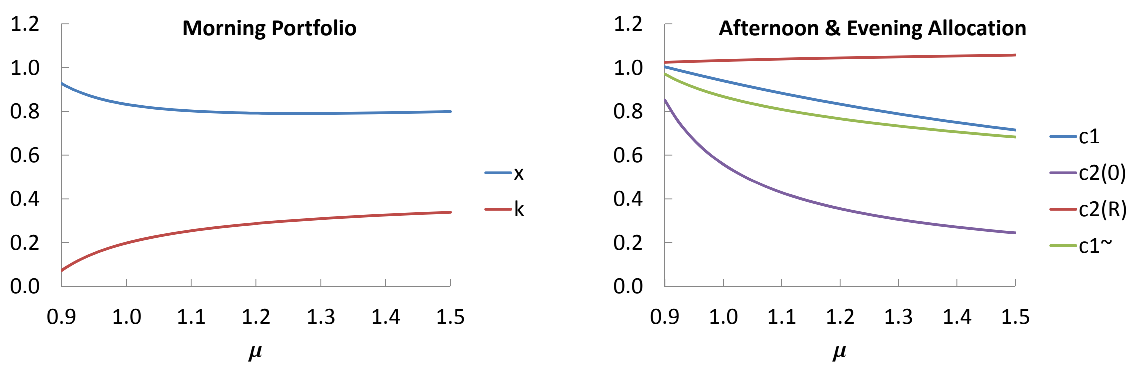


Figure 3: Effects of inflation



(inflation rate financed by lump-sum transfers to workers) (μ). See Figures 1, 2 and 3. The left-hand-side panels in these figures plot the size and composition of bank sector asset positions. The right-hand-side panels plot the consumption allocations.

The left-hand panel in figure 1 shows, unsurprisingly, that bank portfolios are tilted toward cash and away from loans funding capital investments in high θ (inherently fragile) societies. The size of the banking sector's balance sheet, however, seems insensitive to θ. In sufficiently fragile societies, banks hold assets only in the form of cash, refusing to finance any capital spending. The left-hand panel in figure 2 shows that as the expected return to capital spending declines, banks substitute into cash and out of capital. The effect of increasing fundamental uncertainty (and lowering expected capital returns) is similar to the effect of increasing the propensity of coordination failure. The left-hand panel in figure 3 shows how banks substitute away from cash into capital in high inflation environments, reminiscent of a standard Tobin effect.

In terms of the effect of θ on the franchise value of banking, we offer the following proposition.

Proposition 3 *The economic benefit of banking $W^B(\theta, \mu)$ is strictly decreasing in the level of social fragility θ and strictly decreasing in the rate of inflation μ.*

Proof. Differentiating (25) with respect to θ yields

$$dW^B/d\theta = \theta \left\{ u\left(\tilde{c}_1\right) - \pi u(c_1) - (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right]\right\}$$

So, $sign\left[dW^B/d\theta\right] = -sign\left[\pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right] - u\left(\tilde{c}_1\right)\right] < 0$ (we need a more formal proof here). Next, let $\lambda_1, \lambda_2(R_t)$ denote the Lagrange multipliers associated with the constraints (26) and (27). Then differentiating (25) with respect to μ yields

$$dW^B/d\mu = -\mu^{-1}\left[\theta u'\left(\tilde{c}_1\right) + \lambda_1 + \lambda_2(0) + \lambda_2(R)\right] < 0$$

since $\lambda_1 = 0$. ■

7 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. That is, the bank regime objective (25) now reduces to

$$W^N(\mu) \equiv \max\left\{-x - k + \pi u(c_1) + (1-\pi)\left[\delta u(c_2(0)) + (1-\delta)u(c_2(R))\right]\right\} \tag{30}$$

The constraints are now given by

$$\begin{aligned} x\left(\beta/\mu\right) &\geq c_1 &\quad (31)\\ R_t k + \left[x\left(\beta/\mu\right) - \pi c_1\right] &\geq (1-\pi)c_2(R_t) \text{ for } R_t \in \{0, R\} &\quad (32) \end{aligned}$$

Notice that the cash reserve constraint (31) is significantly more stringent than (26). Let λ_1 denote the Lagrange multiplier for (31) and let $\lambda_2(R_t)$ denote the Lagrange multipliers for the state-contingent constraints in (32). Then differentiating with respect to $c_1, c_2(0), c_2(R), x$ and k, respectively,

$$\begin{array}{l} u'(c_1) = \lambda_1/\pi + \lambda_2(0) + \lambda_2(R) \\ \delta u'(c_2(0)) = \lambda_2(0) \\ (1-\delta)\, u'(c_2(R)) = \lambda_2(R) \\ \lambda_1 + \lambda_2(0) + \lambda_2(R) = \mu/\beta \\ \lambda_2(R)R = 1 \end{array} \tag{33}$$

Lemma 4 *The narrow bank solution entails a binding cash reserve constraint* ($\lambda_1 > 0$).

Proof. Imposing $\lambda_1 = 0$ in system (33) implies

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

Note further that $\mu/\beta > [\mu/\beta - 1/R]\,\delta^{-1}$. This implies that $c_1 > c_2(0)$. In any equilibrium, we have $\pi c_1 + (1-\pi)\,c_2(0) = (\beta/\mu)\,x$. The previous inequality yields $c_1 > (\beta/\mu)\,x$, which is a contradiction. ■

Since $\lambda_1 > 0$, we have $c_1 = (\beta/\mu)\,x$. Since the state-contingent budget constraints in (32) bind tightly as before, we can derive

$$\begin{aligned} c_2(0) &= (\beta/\mu)\,x \\ c_2(R) &= \frac{Rk}{1-\pi} + (\beta/\mu)\,x \end{aligned}$$

Combining these latter results with the system in (33) permits us to characterize the optimal bank portfolio under the narrow banking regime

$$\left[\frac{\mu/\beta - (1-\pi)\,/R}{\pi + (1-\pi)\,\delta}\right] = u'((\beta/\mu)\,x) \tag{34}$$

$$(1-\delta)\,Ru'\left(\frac{Rk}{1-\pi} + (\beta/\mu)\,x\right) = 1 \tag{35}$$

Proposition 5 ***Effect of inflation on narrow banking asset composition:*** *$dx/d\mu < x/\mu$ and $dk/d\mu > 0$.*

Proof. Totally differentiate (34) and (35) with respect to μ. We obtain:

$$\begin{aligned} [\beta(\pi + (1+\pi)\delta)]u''(c_1)[(\beta/\mu)(dx/d\mu) - (\beta/\mu^2)x] &= 1 \\ (1-\delta)Ru''(c_2(R))[(R/(1-\pi))(dk/d\mu) + (\beta/\mu)(dx/d\mu) - (\beta/\mu^2)x] &= 0 \end{aligned}$$

Since $u''(c) < 0$, the first expression implies $(\beta/\mu)(dx/d\mu) - (\beta/\mu^2)x < 0$ and so, $dx/d\mu < x/\mu$. From the second expression, we get $(R/(1-\pi))(dk/d\mu) + (\beta/\mu)(dx/d\mu) - (\beta/\mu^2)x = 0$. Therefore, $dk/d\mu > 0$. ■

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 6 *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.*

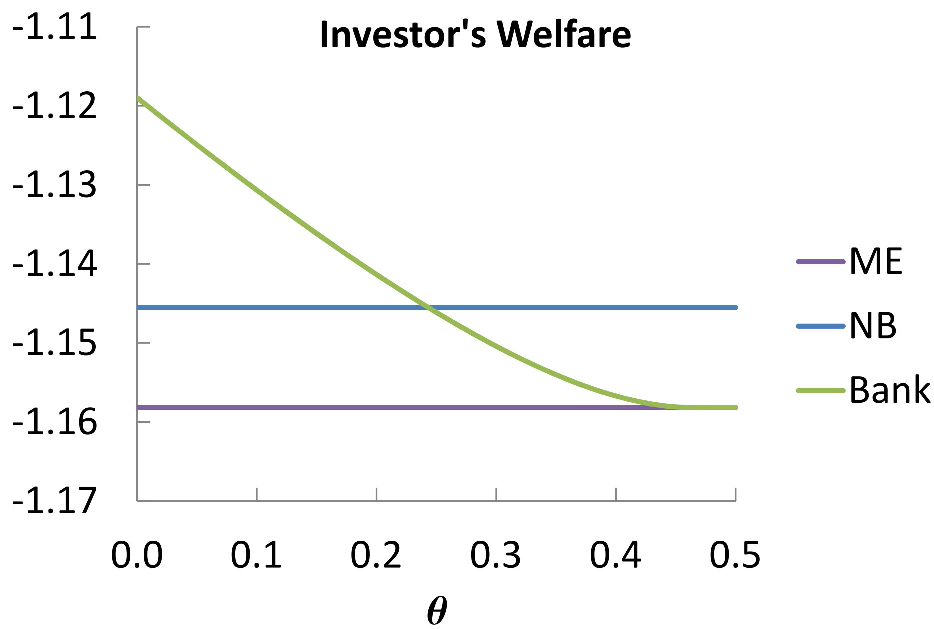
Proof. That $W^N(\mu)$ is independent of θ is obvious. To establish the latter result, differentiate (30) with respect to μ to derive

$$dW^N/d\mu = -\mu^{-1}\,[\lambda_1 + \lambda_2(0) + \lambda_2(R)] < 0$$

■

We conjecture that 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. For the parameter values in (29), we can verify that this is indeed the case (see Figure 4).

Figure 4: Fractional vs Narrow Banking



Note. ME: monetary equilibrium. NB: narrow banking. Bank: fractional banking.

Proposition 7 *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 (25) 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. ■

7.1 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.

7.2 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.

Remark 8 *There is a political economy aspect that we could pursue here. In societies where the median voter is a worker and banking regulation is chosen by the median voter, workers are likely to choose more stable banking systems. So, societies where banks wield less political power, we'd expect fewer financial crises. The Dodd-Frank act could be interpreted as a shift in political power away from banks to other segments of society. Something to explore here.*

8 Lender of last resort facility

This section is preliminary and incomplete.

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.

Our stylized LOLR facility operates as follows. First, 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.[16]

If the facility is open, banks are able to borrow up to the full market value of their collateral.[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. In any case, we assume here 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.[18] In effect, there is a helicopter drop of money to investors when fundamentals turn out to be bad.[19]

The purpose of the LOLR facility, of course, is to prevent inefficient liquidation events. To this end, we'll have to check whether the available collateral $(1 - \delta)Rk$ is sufficient to permit the bank to make good on its short-term obligations. Assume, for the moment, that this is the case. Note that the LOLR must extend its lending prior to knowing the future state of the world, R_t. If it did somehow know R_t at the beginning of the afternoon, an optimal strategy would be to offer the loan in the event $R_t = R$ and decline it in the event $R_t = 0$. In the latter case, the bank is "insolvent" so that early liquidation is efficient–see Allen and Gale (1998). Because our central bank is not so prescient, its lending decision must be made before R_t is known. Consequently, there will be times when our central bank will, on an *ex post* basis, regret having intervened or not intervened. In particular, the following four outcomes are possible following any period in which a sunspot occurs:

	$R_t = 0$ w.p. δ	$R_t = R$ w.p. $(1 - \delta)$
Facility open w.p. α	inefficient intervention	efficient intervention
Facility closed w.p. $1 - \alpha$	efficient non-intervention	inefficient liquidation

Note that the inefficiency associated with the "inefficient intervention" case will be associated with the inflation tax associated with helicopter drops of money that are used to "bail out" failed banks.

[16]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.

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

[18]It would be of some interest to endogenize the repayment choice. For example, the recipient bank might want to repay the central bank in the subsequent morning regardless of how fundamentals turn out in the evening because maintaining a good credit history will maintain access to the LOLR facility.

[19]We could consider alternative ways to finance the LOLR program, but we think money finance is the most natural way to start.

Let us now 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. Upon their arrival, investors discover whether the LOLR facility is open or not. If it is open, the central bank makes a collateralized loan to the private bank, where the loan size is sufficient to finance all short-term obligations c_1. Since the bank holds $x(\beta/\mu) = (1-\pi)c_2(0) + \pi c_1$ units of real cash balances and has promised c_1 consumption to all agents, it requires a real loan of

$$[c_1 - x(\beta/\mu)] = (1-\pi)[c_1 - c_2(0)] = (1-\pi)[c_1 - c_2(R)] + Rk \qquad (36)$$

The private bank then pays out all its cash to investors, fulfilling its short-term obligations without liquidating capital. Next, impatient investors take their cash and spend it on afternoon output (as usual). Patient investors carry their money into the evening.

After having received the cash from the DD-bank, the investors must decide what to do with it. Clearly, the impatient investors spend it in the afternoon to get c_1 consumption. The patient investors have to decide how much to spend in the afternoon and how much to spend in the evening. We assume for now that they replicate the same spending pattern as if there was no run.[20] Since they hold c_1 real units of money, they acquire $c_2(0)$ good in the afternoon and save $c_1 - c_2(0)$ units of real money balances.

If $R_t = R$, they then exchange $c_1 - c_2(0)$ units of money for $Rk(1-\pi)^{-1}$ goods in the evening at the DD-bank. The bank, thus receives $(1-\pi)[c_1 - c_2(0)]$ units of money back, which allows it to pay back the loan. In this case the money supply is unchanged.

If $R_t = 0$, they save and bring $c_1 - c_2(0)$ real units of money into the following period, where it allows them to reduce their labor effort. The bank cannot pay back its loan and defaults. Consequently, the real money supply increases by $(1-\pi)[c_1 - c_2(0)]$.

8.1 LOLR financed by lump-sum tax

There are different ways to model the financing of the LOLR program. The simplest way is to assume that it is financed with a lump-sum tax (that may be applied either on workers or investors, or both). Financing the program in this way implies no inflationary consequences.

As discussed above, when there is a run the DD-bank redeems $p_t^a c_1$ units of money. If the investor is patient, he keeps $p_t^a [c_1 - c_2(0)]$ units of money for the

[20]We will have to provide a formal justification for this behavior but we conjecture that it is optimal.

evening. If $R_t = 0$, he doesn't spend these units and carries them over to the next period morning market. The real value of this quantity is $p_t^a \left[c_1 - c_2(0) \right] / p_{t+1}^m = (p_t^a / p_t^m) \left[c_1 - c_2(0) \right] p_t^m / p_{t+1}^m = (1/\beta) \left[c_1 - c_2(0) \right]$. Because preferences are linear in the morning, an investor values that money with utility $c_1 - c_2(0)$ the next day.

Under these assumptions, the bank's program is as follows:

$$
\begin{aligned}
& -x - k + (1-\theta) \left\{ \pi u(c_1) + (1-\pi) \left[\delta u(c_2(0)) + (1-\delta) u(c_2(R)) \right] \right\} && (37) \\
& +\theta\alpha \left\{ \begin{array}{c} \delta \left[\pi u(c_1) + (1-\pi) \left[u\left(c_2(0)\right) + c_1 - c_2(0) \right] \right] \\ + (1-\delta) \left[\pi u(c_1) + (1-\pi) u(c_2(R)) \right] \end{array} \right\} && (38) \\
& +\theta \left(1-\alpha\right) u \left[(\beta/\mu) x + \xi k \right] && (39)
\end{aligned}
$$

We can derive the first-order conditions as before.

8.2 LOLR financed with inflation tax

It is somewhat more interesting–and realistic, perhaps–to suppose that the LOLR facility is financed with seigniorage. That is, in those cases where banks fail to repay their money loan, the money is not paid back to the central bank and so remains in circulation forever. This manner of financing the LOLR facility has the effect of raising the *expected* rate of inflation–which creates a distortion in the banking sector's optimal portfolio choice.[21]

When there is a run the DD-bank redeems $p_t^a c_1$ units of money, but only holds $p_t^a \left(\beta/\mu\right) x$. It thus has to borrow $p_t^a \left[c_1 - \left(\beta/\mu\right) x \right]$ dollars. The loan is not repaid with probability $\alpha\delta$. Thus, the stock of money evolves probabilistically as follows:

$$
M_{t+1} = \left\{ \begin{array}{lll} \gamma M_t & \text{w/prob} & 1 - \delta\alpha\theta \\ \gamma M_t + p_t^a \left[c_1 - \left(\beta/\mu\right) x \right] & \text{w/prob} & \theta\alpha\delta \end{array} \right.
$$

where γ is the deterministic part of the the money growth rate. Now, μ is the expected inflation rate

$$
\begin{aligned}
\mu &= \gamma + \theta\alpha\delta \left(p_t^a / M_t \right) \left[c_1 - \left(\beta/\mu\right) x \right] \\
&= \gamma + \theta\alpha\delta \left(1-\pi\right) \left[c_1 - c_2(0) \right] \left(p_t^a / M_t \right) \\
&= \gamma + \frac{\theta\alpha\delta \left(1-\pi\right) \left[c_1 - c_2(0) \right]}{\left(\beta/\mu\right) x}
\end{aligned}
$$

and so

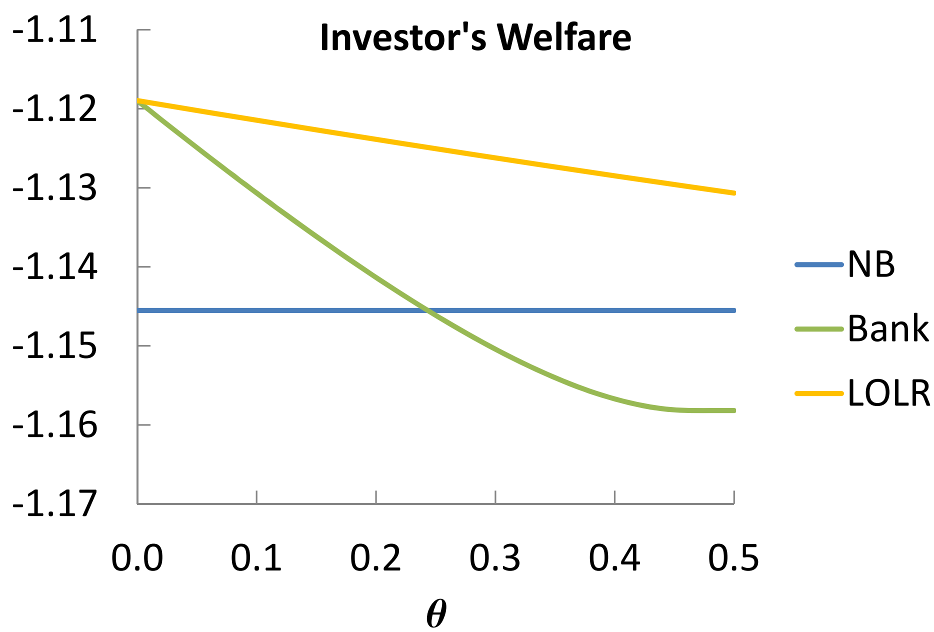
$$
\mu = \gamma + \frac{\theta\alpha\delta \left(1-\pi\right) \left[c_1 - c_2(0) \right]}{\pi c_1 + \left(1-\pi\right) c_2(0)}
$$

[21] Note that expected inflation here will almost always overshoot actual inflation, except in the periods when the bailout actually occurs.

since $p_t^m \mu/\beta = p_t^a$, $p_t^a/M_t = (\mu/\beta)/x$.

Figure 5 shows investors' welfare under an open LOLR facility, as a function of financial fragility θ. The facility is assumed to open 2 out of 3 times there is a sunspot. Naturally, investors benefit from having a bank with access to a LOLR. Perhaps more interesting, the availability of the LOLR facility means that narrow banking may no longer be an attractive alternative. The critical value for θ under which narrow banking is preferred is increasing in α. For $\alpha = 2/3$ narrow banking is never preferred by investors to fractional reserve banking with a LOLR facility.

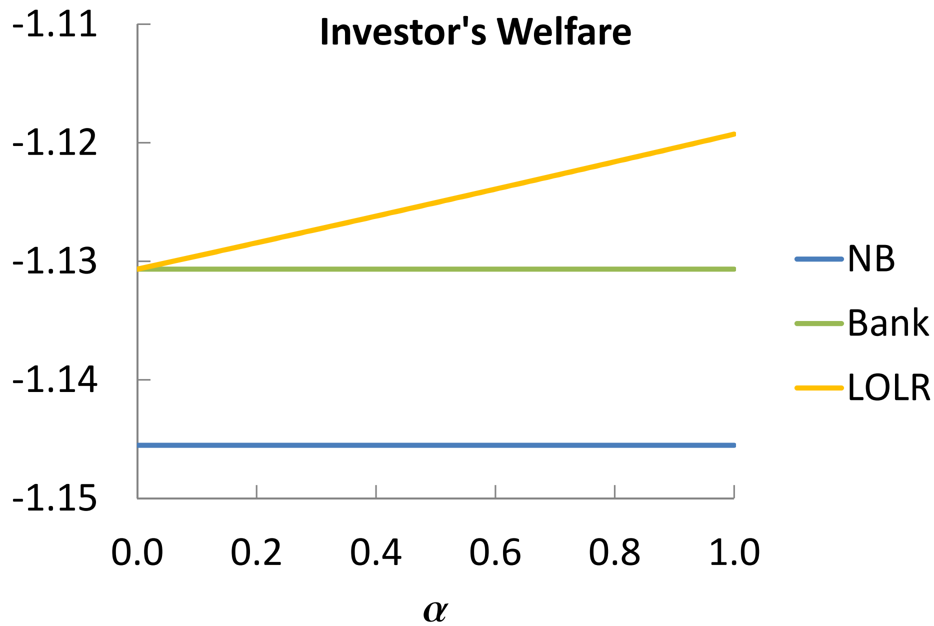
Figure 5: Financial Fragility with LOLR



Note. NB: narrow banking. Bank: fractional banking. LOLR: facility with inflation tax.

Figure 6 shows investor's welfare under an open LOLR facility, as a function of the frequency the facility is open, α. As we can see, welfare is increasing in α, so that it is optimal for investors to set $\alpha = 1$. It is interesting to note that workers' welfare is actually decreasing in α, which is not surprising given that the LOLR facility allows banks to economize on real cash balances. As well, the money injections used to finance the LOLR facility in the bailout events constitutes a transfer of wealth away from workers to bankers. Thus, we again have a disagreement between investors and workers over which banking regime is the preferred arrangement.

Figure 6: Frequency of open LOLR facility



Note. NB: narrow banking. Bank: fractional banking. LOLR: facility with inflation tax.

9 References

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

2. Allen, F., Gale, D., 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, Smith, D. Bruce, 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 Kalin Nikolov (2013). "Government Debt and Banking Fragility: The Spreading of Strategic Uncertainty," NBER Working Paper No. 19278.

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

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

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

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

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

15. 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.

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