Money, Banking and Financial Markets

Authors	David Andolfatto, Aleksander Berentsen, and Fernando M. Martin
Working Paper Number	2017-023F
Revision Date	August 2019
Citable Link	https://doi.org/10.20955/wp.2017.023
Suggested Citation	Andolfatto, D., Berentsen, A., Martin, F.M., 2019; Money, Banking and Financial Markets, Federal Reserve Bank of St. Louis Working Paper 2017-023. URL https://doi.org/10.20955/wp.2017.023

Published In	Review of Economic Studies

Money, Banking and Financial Markets*

David Andolfatto
Federal Reserve Bank of St. Louis

Aleksander Berentsen
University of Basel

Fernando M. Martin
Federal Reserve Bank of St. Louis

August 20, 2019

Abstract

The fact that money, banking, and financial markets interact in important ways seems self-evident. The theoretical nature of this interaction, however, has not been fully explored. To this end, we integrate the Diamond (1997) model of banking and financial markets with the Lagos and Wright (2005) dynamic model of monetary exchange–a union that bears a framework in which fractional reserve banks emerge in equilibrium, where bank assets are funded with liabilities made demandable in government money, where the terms of bank deposit contracts are affected by the liquidity insurance available in financial markets, where banks are subject to runs, and where a central bank has a meaningful role to play, both in terms of inflation policy and as a lender of last resort. Among other things, the model provides a rationale for nominal deposit contracts combined with a central bank lender-of-last-resort facility to promote efficient liquidity insurance and a panic-free banking system.

The views expressed in this paper do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

*We thank the editor, Veronica Guerrieri, three anonymous referees, Zach Bethune, Jeff Campbell, Rodi Manuelli, Guillaume Plantin, Jean-Charles Rochet, Mariana Rojas Breu, Christopher Waller, Steve Williamson, Randy Wright and the participants at several seminars and conferences for many useful comments and criticisms.

1 Introduction

While the fact that money, banking, and financial markets interact in important ways seems sufficiently self-evident, the theoretical nature of this interaction has not been fully explored. Diamond and Dybvig (1983), for example, explain the existence of banking, but do so in a model without money or financial markets. Diamond (1997) explains how banks and financial markets compete as mechanisms for liquidity insurance, but does so in a model without money. Loewy (1991) develops a model with money and banking, but abstracts from monetary policy and financial markets. Smith (2002) presents a model of money, banking, and monetary policy, but abstracts from financial markets. As far as we know, there has been no comprehensive theoretical analysis of how these three phenomena interact with each other and what this might mean for policy.

In this paper, we combine the Diamond (1997) model of banking and financial markets with the Lagos and Wright (2005) model of monetary exchange. The result is a model where fractional reserve banks emerge in equilibrium, where bank assets are funded with liabilities made demandable in government money, where the terms of bank deposit contracts are affected by the liquidity insurance available in financial markets, where banks are subject to runs, and where a central bank has a meaningful role to play, both in terms of inflation policy and as a lender of last resort.

Money, in our model, takes the form of zero-interest nominal government debt. The real rate of return on money–the inverse of the inflation rate–is determined by policy and is financed with lump-sum taxes or transfers. Money is necessary in the economy because an absence of trust between some trading parties precludes the use of credit (Gale, 1978). Money, however, is dominated in rate of return by securities representing claims against an income-generating capital good. Securities are illiquid in the sense they cannot be used to buy consumption goods. But securities possess a degree of indirect liquidity to the extent they can be readily exchanged for money on short notice (Geromichalos and Herrenbrueck, 2019). A financial market in which securities trade for money provides one mechanism for investors to access liquid funds. A bank that stands ready to convert deposit liabilities for cash provides another such mechanism.

Individuals in our model are subject to random liquidity needs as in Diamond and Dybvig (1983). In some periods, people have a pressing need to consume–they are *impatient.* In other periods, they are willing to defer consumption–they are *patient.* Following Allen and Gale (2007, § 3.2), our investigation begins by asking how the economy might function in the absence of banks, but when individuals have access to a market where they can liquidate securities. A well-known conclusion in this body of literature is that the resulting competitive equilibrium is inefficient, except for a knife-edge case relating to the nature of preferences (see also Farhi, Golosov and Tsyvinski, 2009). This conclusion is shown below to be an artifact of the static nature of the models employed. Our first result demonstrates that if monetary policy follows the Friedman rule–and if investors have perfect access to securities markets–then the competitive equilibrium is efficient and that, moreover, the ability to save money across time is critical for this to be true. [1] When securities market

[1] The Friedman rule refers to a policy that eliminates the scarcity of money. The effect of a Friedman

participation is limited as in Diamond (1997), the equilibrium securities price is too low and even the Friedman rule cannot implement the efficient allocation. The economic rationale for a banking system in this environment must therefore stem from one of two frictions, either monetary policy departs from the Friedman rule and/or securities markets are unavailable or sufficiently palsied.

Next, we follow Diamond and Dybvig (1983) and examine how the economy might function in the absence of a securities market, so that cash and securities are held indirectly as bank deposit liabilities. The optimal risk-sharing arrangement entails the use of demandable debt, except that in our case, this debt is made redeemable for government money (instead of goods). We find that a competitive (or monopolistic, but contestable) banking system is also consistent with efficiency, but once again, only when monetary policy follows the Friedman rule. The comparative performance of banks *vis-à-vis* full-participation securities markets is similar at low rates of inflation and is identical when inflation policy is set optimally.

Away from the Friedman rule, the return on money is too low and liquidity is scarce in full-participation financial markets and banking systems alike. However, the impact of inflation in these two systems is not identical. We demonstrate that at higher rates of inflation, the liquidity insurance provided by banks is superior to that provided by even full-participation financial markets. To achieve a similar degree of risk-sharing through securities markets, investors would have to carry precautionary money balances across time– something that is costly to do when the inflation rate is high. In contrast, banks can pool idiosyncratic liquidity risk and so, avoid carrying precautionary money balances across time. Thus, while welfare under both types of arrangements is decreasing in the rate of inflation, banking becomes relatively more valuable at higher inflation rates. Banks also dominate market mechanisms to the extent that securities markets are subject to limited participation and to the extent that banks serve as easy-to-access standing facilities.

The analysis then proceeds to examine the implications of coexistence. Diamond and Dybvig (1983) and many papers that follow in their tradition assume that depositors are prevented from engaging in financial transactions outside of their banking relationships. Bencivenga and Smith (1991) suggest that this is approximately true in developing economies where government regulations often serve to restrict financial markets in favor of banks. But securities markets are relatively well-developed in more advanced economies so that depositors have more options. It has been known since at least Jacklin (1987) that if depositors cannot be refrained from engaging in *ex post* financial market trades, the *ex ante* superior liquidity insurance made possible through banks may not emerge in equilibrium. Indeed, we reproduce the Jacklin (1987) result below. If depositors have free access to a competitive securities market, banks are essentially constrained to offer a liquidity insurance contract that replicates what depositors can achieve on their own in a financial market. Since, as argued above, banking becomes relatively more valuable as inflation increases, the welfare

rule policy is to increase the real rate of return on money to its socially desirable level (in a broad class of models, this corresponds to the rate of time preference). The desired return on money is typically financed with a lump-sum tax. In the case of zero-interest money, the tax is used to finance a deflation. In the case of interest-bearing money, the tax is used to finance the interest expense of money; see Andolfatto (2010) and Grochulski and Zhang (2017).

loss from this "excess competition" is also increasing in the rate of inflation.

Even in more advanced economies, however, access to securities markets is not costless and participation is limited. To examine the implications of coexistence, we take the general equilibrium of our securities market model and interpret it as an unbanked sector of the economy.[2] As already mentioned, when market participation is limited, the equilibrium securities price is too low at *any* inflation rate. We then permit a subset of the population to enter into a banking arrangement. A bank generally dominates what individuals can attain outside of it, with the availability of *ex post* securities trades influencing the properties of the optimal bank contract. While the gains to banking usually depend on the bank discouraging early redemptions by patient depositors, we find that this is no longer the case here. In particular, the under-priced security represents a buying opportunity for patient depositors who have access to the securities market and it makes sense for banks to let their clients exploit it.

When markets and banks coexist, we show that inflation generally affects depositors adversely. Thus, the Friedman rule is still the best policy, though it typically does not implement the first-best allocation—this would require full-participation in financial markets. We also find that the measure of patient depositors accessing financial markets has ambiguous effects on depositor welfare. On the one hand, a larger fraction of depositors accessing financial markets impinges on a bank's ability to provide liquidity insurance. On the other hand, more depositors are able to exploit a profitable investment opportunity, by converting their deposits into securities. Which effect dominates depends on the securities price.

The results reported to this point rest on the assumption that banks are not subject to runs–that is, events in which banks fall into insolvency owing to a fear of insolvency that becomes a self-fulfilling prophecy. In fact, the Diamond and Dybvig (1983) model was motivated by the question of whether the banking system is prone to runs and whether a government deposit insurance scheme could be designed to eliminate them. However, as far as we know, the question of how the presence of a securities market interacts with the phenomenon of bank runs in the Diamond and Dybvig (1983) framework has not been investigated. Our results here are summarized as follows.

As alluded to above, banking arrangements provide superior liquidity risk-sharing outcomes when market-access is limited. But banks are inherently fragile here in the sense of Diamond and Dybvig (1983). In a stress event, illiquid securities may have to be converted to cash. If market access is limited, it will be difficult to liquidate (or otherwise spend) these securities at a fair price. Furthermore, early liquidation coupled with limited market access will prevent some investors from making profitable investments. *Ex post*, the ability to buy and sell securities is enhanced with greater access to markets. But *ex ante*, greater market-access may constrain the ability of a bank to provide superior liquidity insurance. The benefits provided by markets in the event of a bank run are shown to depend on how the bankruptcy is resolved. Our model suggests that dispersing assets in the form of cash and "clearinghouse certificates", similar to those issued by bank coalitions during the panics

[2]Note that "unbanked" here does not necessarily mean "financially unsophisticated." It could, in fact, mean the opposite–i.e., investors with ready access to liquid securities markets have less of a need for banks.

of the U.S. National Banking era, generally dominates outright asset liquidation.

Thus, we identify a potential trade-off: innovations designed to improve securities market liquidity may affect deposit contracts adversely, but conditional on a bank run, lead to better *ex post* liquidation outcomes. Conversely, restrictions on the trading of securities could improve liquidity insurance provided by banks, but leave the economy vulnerable to the dislocations associated with banking runs. This latter prediction is broadly consistent with evidence showing that financial crises tend to be significantly more disruptive in developing economies relative to economies with more developed financial markets; see Reinhart and Rogoff (2009, Figure 4; 2014, Table 2).

The trade-off between insurance and stability vanishes under an appropriate and more comprehensive monetary policy. The source of instability in the Diamond and Dybvig (1983) model is a contractual incompleteness that renders bank deposit liabilities "run-prone".[3] In principle, a fiscal policy that insures deposits against such events could prevent runs from occurring. But the effectiveness of an intervention depends on its credibility and a fiscal intervention must ultimately resort to direct taxation. Relative to fiscal policy, monetary policy has a distinct advantage because the object under its control (cash) also happens to be the object of redemption in demand deposit contracts. Thus, if deposit liabilities are purposely designed to be claims against cash (instead of goods), the monetary authority is always in a position to print the cash necessary to help banks fulfill their nominal obligations. The fact that cash can be printed costlessly enhances the credibility of the intervention and, in this way, the *threat* of such an intervention is sufficiently credible to discourage bank runs.

The paper is organized as follows. In Section 2, we describe the physical properties of the model economy and characterize the nature of an efficient allocation. In Section 3, we introduce the frictions that motivate monetary exchange and we characterize the equilibrium for two separate arrangements, a securities market and a banking system. In Section 4, we examine the implications of the coexistence of these two arrangements and discuss related literature. In Section 5, we examine the consequences of risk-sharing and bank runs against various degrees of securities market liquidity. We also discuss the merits of a central bank lending facility and the desirability of nominal debt. Section 6 concludes. In what follows, we state propositions in the body of the text but relegate all formal proofs to the appendix.

2 The environment

Time, denoted t, is discrete and the horizon is infinite, $t = 0, 1, 2, \ldots, \infty$. Each time period t is divided into three subperiods: *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 (x) and capital (k) goods, so that $y_0 = x + k$. Investors are subject

[3] In particular, if promised redemption rates are made invariant to the volume of early withdrawals, a wave of heavy redemption activity may leave a bank (or the banking system) without enough cash to fulfill its obligations. If the fear of such an event leads depositors to withdraw cash *en masse*, the result is a self-fulfilling bank panic.

to an idiosyncratic preference shock, realized at the beginning of the afternoon, 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 $1 - \pi$. 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 $u'(0) = \infty$. In fact, as is standard in this literature, we assume that investors are sufficiently risk-averse in the following sense:

[A1] $-u''(c)c > u'(c)$ for all $c \geq 0$.

Investors discount utility payoffs across periods with subjective discount factor $0 < \beta < 1$, so that investor preferences are given by

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

Workers have linear preferences defined over 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 of the same period, but are perishable across periods. Workers share the same discount factor as investors, so that worker preferences are given by

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

The investment technology available to investors works as in the manner described in Cooper and Ross (1998). In particular, capital goods produced in the morning (k) are assumed to generate a high rate of return ($R > 1$) if left to gestate into the evening and a low rate of return ($0 < r < 1$) if gestation is interrupted and operated in the afternoon.[4] For simplicity, we assume that capital depreciates fully after it is used in evening production.[5]

To characterize 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 a given number (normalized here to zero, so that $c_0 = y_1$). The condition $r < 1$ implies that workers can supply afternoon output more efficiently than investors who operate capital prematurely in the afternoon. As a consequence, efficiency dictates that scrapping capital in the afternoon is never optimal. In addition, the morning resource constraint $x = c_0$ must

[4] In nonmonetary models, the parameter r is sometimes described interchangeably as the "scrap" or "liquidation" value of capital. In our monetary model, however, r represents the consumption of capital, not its liquidation value. In other words, capital may be either scrapped (consumed) or exchanged for money (liquidated).

[5] We chose this specification in part for its analytical tractability and in part because the model nests the standard Diamond and Dybvig (1983) model when $\beta = 0$. We do not think our main conclusions are sensitive to the existence of a durable asset, though this is something that remains to be verified.

hold since the morning transfer of utility from investors to workers must balance. With these conditions imposed, the planning problem may be written as

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

subject to

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

The *first-best allocation* is characterized by the following two restrictions,

$$u'(c_1^*) = Ru'(c_2^*) = 1 \tag{6}$$

where, using (4) and (5), $x^* = \pi c_1^*$ and $k^* = (1-\pi)c_2^*/R$. For the special case $u'(c) = c^{-\sigma}$, $\sigma > 1$, we have,

$$c_1^* = 1 \text{ and } c_2^* = R^{1/\sigma} \tag{7}$$

If, in addition, investor types are private information, then incentive-compatibility conditions are necessary to ensure truthful revelation. Incentive-compatibility requires $c_2 \geq c_1$ for patient investors and $c_1 \geq 0$ for impatient investors. Note that (6) implies that $c_2^* > c_1^* > 0$, so that the first-best allocation is incentive-compatible. Note too that conditions (6) imply $u'(c_1^*) = Ru'(c_2^*)$ as in the standard Diamond-Dybvig model. However, unlike the standard model, ours is a production economy with linear costs, which is why both $u'(c_1^*)$ and $Ru'(c_2^*)$ need to equal the constant marginal cost of production at the first-best.

3 Monetary equilibrium with a securities market

Workers and investors meet in the morning and in the afternoon (but not in the evening). Ideally, investors would like to borrow output from workers in the afternoon, repaying the loan in the next morning. Such credit arrangements are ruled out here because workers and investors are assumed not to trust each other (Gale, 1978). As a consequence, any trade between workers and investors must occur on a *quid-pro-quo* basis through the use of an exchange medium. The exchange medium here is assumed to take the form of a zero-interest-bearing government debt instrument (money), the total nominal 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 (this initial condition is immaterial for what follows). New money is created at the beginning of each morning at the constant rate μ, so that $M_t = \mu M_{t-1}$. We assume that new money $T_t = (1 - 1/\mu) M_t$ is injected as lump-sum transfers to workers.[6] Because preferences are linear in morning consumption, the positive implications of our analysis are unaffected by precisely who gets the new money although, of course, there are redistributional consequences. We

[6]While we permit any amount of deflation here in the range $\beta \leq \mu < 1$, there is the question of whether lump-sum taxation is an incentive-feasible policy. Andolfatto (2013) addresses this issue, but we ignore it in what follows.

assume that trade in goods (for money) occurs on a sequence of competitive spot markets. Let (p_t^m, p_t^a) denote the dollar price of output in the morning and afternoon, respectively.

Workers produce output in the afternoon for money, and carry the money forward to the next morning where they spend it on consumption. Because workers have linear preferences, their cost-benefit calculation is simple to characterize. To acquire a dollar in the afternoon, a worker must expend $1/p_t^a$ units of labor. This dollar is convertible for $1/p_{t+1}^m$ units of output the following morning. In present value terms, this future output is worth β/p_{t+1}^m units of afternoon t output. For an arbitrary price-system, the solution is at one of two corners: workers will want to work as much as possible, or not at all. In equilibrium, the output prices adjust so that workers are indifferent between working and not working. Hence,

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

Condition (8) implies that workers are willing to supply output passively (elastically) in the afternoon to accommodate any level of investor demand.

We focus our analysis on stationary monetary equilibria. As is well known for this class of models (e.g., see Lagos and Wright, 2005), in a stationary monetary equilibrium, nominal prices grow at the same rate as the money supply. That is,

$$p_{t+1}^m/p_t^m = p_{t+1}^a/p_t^a = \mu \tag{9}$$

Combining this latter restriction with condition (8) implies,

$$p_t^m/p_t^a = \beta/\mu \tag{10}$$

Condition (10) says that when monetary policy is away from the Friedman rule (i.e., $\beta < \mu$), the price of afternoon money is discounted relative to morning money, i.e., $1/p_t^a < 1/p_t^m$. This is because money acquired in the afternoon can only be spent the next day and, away from the Friedman rule, money earns less than its socially optimal rate of return (i.e., $1/\mu < 1/\beta$).[7]

Condition (10) expresses the real rate of return to money held from the morning to afternoon. The return to capital goods held over the same period is given by the parameter r. In what follows, we assume that $\beta/\mu > r$ so that money is preferred to capital goods as a "short-run" investment vehicle. We also assume that $\mu \geq \beta$, since monetary equilibria otherwise fail to exist. To summarize, throughout the remainder of the paper we assume (with the Friedman rule $\mu = \beta$ understood to mean the limiting case as $\mu \searrow \beta$):

[A2] $r < \beta/\mu < 1$.

Because $R > 1$, capital goods (or claims to capitals goods, which below we refer to as *securities*) are preferred to money as a "long-run" investment vehicle. But the willingness of investors to hold securities will depend in part on how easily they can be liquidated (convertible to cash) in the random event they become impatient. This, in turn, depends

[7] Normally, one solves for individual decisions for an arbitrary price sequence. In what follows, we impose condition (10) on the choice problem faced by investors, anticipating that we can do so here without harm.

on the availability of a securities market, or some institution (like a bank) that stands ready to perform the conversion. In what follows, we study the role of markets and banks as alternative mechanisms for providing the desired liquidity insurance.

Following Allen and Gale (2007, § 3.2), we first examine how this economy might solve the "liquidity problem" through a competitive market in which investors exchange securities for money in the afternoon after their preference types have been realized. In fact, in what follows, we adopt the more general market structure introduced by Diamond (1997), who assumes an exogenous fraction $\eta \in [0,1]$ of patient investors are randomly chosen to participate in the afternoon securities market. Impatient investors are all assumed to have market access. Investors realize their market access shock together with the realization of their preference shock. This specification conveniently nests the models of Diamond and Dybvig (1983) and Allen and Gale (2007, § 3.2) as special cases; i.e., when $\eta = 0$ and $\eta = 1$, respectively.

The analysis that follows is simplified by the fact that, in equilibrium, workers choose not to participate in the securities market (see Appendix B). Given this result, we focus on investor behavior in what follows.

The wealth portfolios investors carry with them from the end of the evening to the following morning will generally depend on their afternoon trading histories. Let $m_t^m \geq 0$ denote the money balances brought into the morning of period t by an investor. Note that since capital goods depreciate fully at the end of each period, all investors will enter the morning with zero securities. Define real money balances at the beginning of the morning as $z_t \equiv m_t^m / p_t^m$ and at the end of the morning by $x_t \equiv m_t^a / p_t^m$.

In the morning, investors will want to rebuild their depleted wealth portfolios. They can do so by producing morning output y_0 and then selling some of it in the form of consumer goods $x_t - z_t$ to workers at price p_t^m for cash m_t^a and by storing the remainder as capital goods k_t (investment). Because investor preferences are linear in morning output, all investors choose the same end-of-morning wealth portfolio (x_t, k_t), regardless of their potentially different beginning-of-morning wealth position z_t.

In the afternoon, investors learn whether they are patient or impatient and, if patient, whether they have access to the securities market or not. Impatient investors will want to sell securities and patient investors will want to buy them. Let $k_{1,t}$ and $k_{2,t}$ denote the securities sold and bought, by an impatient and patient investor with market access, respectively. Let q_t denote the nominal price of a security (a claim to a unit of the capital good, measured in units of money). An impatient investor faces the constraint $k_t \geq k_{1,t}$ while a patient agent with market access faces the constraint $m_t^a \geq q_t k_{2,t}$.

Because $x_t = m_t^a / p_t^m$, the real value of money balances in the afternoon is given by $m_t^a / p_t^a = (p_t^m / p_t^a) x_t$. Using condition (10), we can write $m_t^a / p_t^a = (\beta/\mu) x_t$. It will also prove useful to define $\rho_t \equiv q_t / p_t^a$, the afternoon price of securities measured in units of afternoon output. All investors carry their morning wealth $x_t + k_t$ into the afternoon, where it is worth $(\beta/\mu) x_t + \rho_t k_t$, measured in units of afternoon output. Let $m_{1,t+1}^m, m_{2,t+1}^m, \hat{m}_{2,t+1}^m \geq 0$ denote the money carried forward to the next morning for the impatient and patient investors (with and without market access at t), respectively. Using (8) and the fact that (in a stationary equilibrium) the equilibrium inflation rate is given by

$$p_{t+1}^m / p_t^m = \mu \tag{11}$$

we can derive $m_{1,t+1}^m / p_t^a = \beta z_{1,t+1}$, $m_{2,t+1}^m / p_t^a = \beta z_{2,t+1}$,and $\hat{m}_{2,t+1}^m / p_t^a = \beta \hat{z}_{2,t+1}$.

Now consider the budget equation for an impatient investor, $c_{1,t} + m_{1,t+1}^m / p_t^a = m_t^a / p_t^a + (q_t / p_t^a) k_{1,t} + r(k_t - k_{1,t})$. Using the notation described above, rewrite this equation as

$$c_1 = (\beta/\mu)x + \rho k_1 + r(k - k_1) - \beta z_1^+ \tag{12}$$

where here we have suppressed time subscripts (adopting the convention $z_{t+1} = z^+$) to ease notation. The impatient investor is subject to the following constraints,

$$\begin{aligned} z_1^+ &\geq 0 && (13) \\ k - k_1 &\geq 0 && (14) \\ (\beta/\mu)x + \rho k_1 - \beta z_1^+ &\geq 0 && (15) \end{aligned}$$

Condition (13) says that cash balances carried forward in time cannot be negative (there is no borrowing). Condition (14) says that the impatient investor cannot sell more securities in the afternoon market than he brings into the period. Condition (15) says that the amount of cash brought into the next period cannot exceed the amount of cash brought into the afternoon augmented by the cash acquired via afternoon sales of securities.

Next, consider the budget equation for a patient investor with market access, $c_{2,t} + m_{2,t+1}^m / p_t^a = R(k_t + k_{2,t}) + m_t^a / p_t^a - (q_t / p_t^a) k_{2,t}$. Again, using the notation described above, rewrite this budget equation as

$$c_2 = R(k + k_2) + (\beta/\mu)x - \rho k_2 - \beta z_2^+ \tag{16}$$

The patient investor with market access is subject to the following constraints,

$$\begin{aligned} z_2^+ &\geq 0 && (17) \\ (\beta/\mu)x - \rho k_2 &\geq 0 && (18) \\ (\beta/\mu)x - \rho k_2 - \beta z_2^+ &\geq 0 && (19) \end{aligned}$$

Condition (17) stipulates that cash balances going forward must be non-negative. Condition (18) states that the value of securities purchased in the afternoon market cannot exceed the amount of cash on hand. Condition (19) says that the amount of cash carried forward plus cash spent on securities cannot exceed the amount of cash on hand.

Finally, consider a patient investor that has no access to the securities market. His budget constraint is $\hat{c}_{2,t} + \hat{m}_{2,t+1}^m / p_t^a = R k_t + m_t^a / p_t^a$, which can be rewritten as:

$$\hat{c}_2 = (\beta/\mu)x + Rk - \beta \hat{z}_2^+ \tag{20}$$

The patient investor without market access is subject to the following constraints,

$$\begin{aligned} \hat{z}_2^+ &\geq 0 && (21) \\ (\beta/\mu)x - \beta \hat{z}_2^+ &\geq 0 && (22) \end{aligned}$$

Conditions (21) and (22) state that cash balances going forward must be non-negative and cannot exceed the amount of cash on hand.

Let $V(z)$ denote the value of an investor who enters the morning with real money balances z. This value function must satisfy the following recursive relationship:

$$V(z) \equiv \max_{x,k,k_1,k_2,z_1^+,z_2^+,\hat{z}_2^+} z - x - k + \pi\left[u(c_1) + \beta V(z_1^+)\right] + (1-\pi)\eta\left[u(c_2) + \beta V(z_2^+)\right] + (1-\pi)(1-\eta)\left[u(\hat{c}_2) + \beta V(\hat{z}_2^+)\right] \quad (23)$$

subject to (12)–(22).

We now provide some important properties of a securities market equilibrium. We characterize general properties of the equilibrium securities price and how they depend on η. In particular, we show that both money and capital are held in positive amounts by investors and establish that constraints (15) and (18) are typically redundant, while constraints (14) and (19) bind.[8] The first-order conditions characterizing optimal investor behavior are reported in the proof.

Proposition 1 *Properties of a securities market equilibrium: (i) for* $\eta \in (0,1)$*, the equilibrium securities price satisfies* $0 < \rho < \beta/\mu$ *and when* $\eta = 1$*, the equilibrium securities price is given by* $\rho = \beta/\mu$*; (ii)* $x > 0$ *and* $k > 0$*; (iii) when* $r < \rho \leq \beta/\mu$ *constraints* (15) *and* (18) *are slack, while constraints* (14) *and* (19) *bind.*

We now characterize the securities market equilibrium focusing on the case $r < \rho \leq \beta/\mu$. Consider the choice of x in the morning, which can be thought of as the demand for real money balances. Adding one unit of real money balances to the morning wealth portfolio costs the investor one unit of morning output (or utility) and is worth (β/μ) units of afternoon output; see (10). What is the expected benefit of this additional money? With probability π the investor turns out to be impatient in the afternoon. In this case he consumes (β/μ) in the afternoon–the expected marginal benefit is given by $\pi(\beta/\mu)u'(c_1)$. With probability $(1-\pi)\eta$ the investor turns out to be patient with market access. Market access means the investor can use his afternoon real money balances (β/μ) to buy securities that pay off $(R/\rho)(\beta/\mu)$ units of evening consumption–the expected marginal benefit is given by $(1-\pi)\eta(R/\rho)(\beta/\mu)u'(c_2)$. With probability $(1-\pi)(1-\eta)$ the investor turns out to be patient without market access. In this case, the investor can use his afternoon real money balances to buy (β/μ) units of afternoon consumption and store it into the evening, and/or to save money balances for the following period–the expected marginal benefit is given by $(1-\pi)(1-\eta)(\beta/\mu)[u'(\hat{c}_2) + \hat{\lambda}_2]$, where $(1-\pi)(1-\eta)\hat{\lambda}_2$ is the Lagrange multiplier associated with constraint (22). Collecting these terms, we have the first-order condition,

$$1 = (\beta/\mu)\{\pi u'(c_1) + (1-\pi)\eta(R/\rho)u'(c_2) + (1-\pi)(1-\eta)[u'(\hat{c}_2) + \hat{\lambda}_2]\} \quad (24)$$

[8] As we show in Appendix A, these properties hold when $r < \rho$, which is the case we focus on below. Constraint (15) binds when $0 < \rho < r$, while constraint (14) is slack when $0 < \rho \leq r$. Note that in static environments, $r \leq \rho$ as impatient investors would never sell securities below their scrap value. In our dynamic environment, when the securities price is very low ($\rho < r$), it may be possible that impatient investors find it optimal to instead buy (and then scrap) securities from patient investors, who in turn, would sell them to accumulate cash balances, which they would carry to the following period to save on effort.

Consider next the choice of k in the morning, which corresponds to the demand for investment. Adding one unit of capital goods to the morning wealth portfolio costs the investor one unit of morning output (or utility). This unit of capital can be scrapped in the afternoon for r units of output, but there will generally be a better alternative available. In particular, since impatient investors have access to the securities market, they could securitize this unit of capital and sell it for $\rho > r$ units of afternoon output–the expected marginal benefit is given by $\pi\rho u'(c_1)$. With probability $(1-\pi)$ the investor turns out to be patient, in which case he holds on to the capital for a return equal to R in the evening. This additional level of evening output is evaluated differently at the margin depending on whether the patient investor has market access or not. If he has market access, then the expected marginal benefit is given by $(1-\pi)Ru'(c_2)$. If he does not have market access, then the expected marginal benefit is given by $(1-\pi)Ru'(\hat{c}_2)$. Collecting these terms, we have the first-order condition,

$$1 = \rho\pi u'(c_1) + (1-\pi)\eta R u'(c_2) + (1-\pi)(1-\eta)Ru'(\hat{c}_2) \tag{25}$$

Combining (24) with (25), we can derive the following expression

$$1/\rho - \mu/\beta = (1-\pi)(1-\eta)[(R/\rho - 1)u'(\hat{c}_2) - \hat{\lambda}_2] \tag{26}$$

When patient investors without market access spend some cash on afternoon goods, constraint (22) is slack, i.e., $(1-\eta)\hat{\lambda}_2 = 0$. Then, (26) determines the level of consumption $\hat{c}_2$ as a function of ρ and parameters. As we show in the proof of Proposition 1, when (22) instead binds, $0 < \hat{\lambda}_2 < (R/\rho - 1)u'(\hat{c}_2)$. In this case, patient investors without market access only consume the return on their capital investment and save all their money holdings, i.e., $\hat{c}_2 = Rk$, $\beta\hat{z}_2^+ = (\beta/\mu)x$ and $\hat{\zeta}_2 = 0$.

The remaining constraints to consider concern the non-negativity conditions on cash carried from one time period to the next; i.e., (13), (17) and (21). These constraints either bind or not. The complementary slackness conditions are as follows.

$$\begin{aligned} [u'(c_1) - 1]z_1^+ &= 0 && (27) \\ \eta[(R/\rho)u'(c_2) - 1]z_2^+ &= 0 && (28) \\ (1-\eta)[u'(\hat{c}_2) - 1 + \hat{\lambda}_2]\hat{z}_2^+ &= 0 && (29) \end{aligned}$$

These conditions imply upper bounds on consumption, in particular, $u'(c_1) \geq 1$, $u'(c_2) \geq \rho/R$ and $u'(\hat{c}_2) \geq 1 - \hat{\lambda}_2$. Note that when $\hat{\lambda}_2 > 0$ we obtain $u'(\hat{c}_2) > \rho/R$.

The amount of securities sold in equilibrium must equal the amount purchased, so that

$$\pi k_1 = (1-\pi)\,\eta k_2 \tag{30}$$

By Proposition 1 (and the fact we are restricting attention to $r < \rho$) an impatient investor wants to sell all of his securities $k_1 = k$ and a patient investor with market access is willing (and able) to acquire all securities offered for sale (carrying any leftover money across time).

We next characterize the consumption level and quantity of money investors carry with them from one period to the next. Consider first those patient investors with market access.

Combining the fact that $k_1 = k$ with (30), implies

$$k_2 = \left[\frac{\pi}{(1-\pi)\eta}\right] k \tag{31}$$

Because condition (19) binds, we have $(\beta/\mu)x = \rho k_2 + \beta z_2^+$ which, when combined with (16) implies $c_2 = R(k + k_2)$. Combining this latter expression with (31) implies,

$$c_2 = \left[\frac{\pi + (1-\pi)\eta}{(1-\pi)\eta}\right] Rk \tag{32}$$

Note that the level of consumption enjoyed by the patient investor with market access c_2 depends on the price of securities only through any indirect effect ρ might have on the morning investment decision k. Note too that c_2 is financed entirely out of the returns to capital invested in the morning k. Recall that it is feasible for investors to purchase afternoon output and store it for evening consumption. But for investors with market access, this latter option is never taken–it is more desirable to spend money on securities and carry any excess money over to the next period. Again, recall that constraint (19) binds, so that $(\beta/\mu)x - \rho k_2 = \beta z_2^+$. Combining this latter expression with (31) implies,

$$\beta z_2^+ = (\beta/\mu)x - \left[\frac{\pi}{(1-\pi)\eta}\right] \rho k \geq 0 \tag{33}$$

The evolution of future cash balances for impatient investors and investors without market access are given, respectively, by

$$\beta z_1^+ = (\beta/\mu)x + \rho k - c_1 \quad \geq \quad 0 \tag{34}$$
$$\beta \hat{z}_2^+ = (\beta/\mu)x + Rk - \hat{c}_2 \quad \geq \quad 0 \tag{35}$$

A sharper characterization of the equilibrium requires identifying regions in the parameter space where various combinations of the inequality constraints (13), (17), (21) and (22) bind. When (17) binds, $z_2^+ = 0$, and so, from (33) the equilibrium securities price will satisfy,

$$\rho = \eta \left(\frac{\beta}{\mu}\right)\left(\frac{1-\pi}{\pi}\right)\left(\frac{x}{k}\right) \tag{36}$$

Condition (36) shows clearly how the parameters (η, μ) have both a direct and indirect effect on the equilibrium securities price ρ. The direct effect of an increase in η is to increase the price of securities. The intuition for this is that there are now more patient investors with cash able to buy securities. A change in the market-access probability will, in addition, alter the portfolio choice of investors; higher η is likely to translate into a higher cash-securities ratio, (x/k), and so pressure ρ upwards as well, since it becomes more likely to be able to dispose of cash to buy securities when patient. A higher securities price, however, is likely to spur morning investment k at the expense of cash holdings, so (x/k) is likely to decrease. We cannot say at this level of generality whether this latter effect will be outweighed by the former.

The direct effect of an increase in μ is to decrease the price of securities. The intuition is that, by arbitrage, a lower short-term rate of return on cash, β/μ implies a lower short-term rate of return on capital, ρ. The effect of μ on the investor's portfolio is ambiguous and depends on preferences. A lower securities price, however, is likely to depress morning investment k at the expense of cash holdings, so (x/k) is likely to increase. Again, we cannot say at this level of generality whether this latter effect will be outweighed by the former, except for when $\eta = 1$, in which case $\rho = \beta/\mu$.

3.1 Full participation

In this section, we consider the Allen and Gale (2007, § 3.2) case in which $\eta = 1$. Proposition 1 implies that the equilibrium securities price ρ for the case of $\eta = 1$ depends directly on parameters, one of which is the money supply growth (inflation) rate. The condition $\rho = \beta/\mu$ can be usefully thought of here as a no-arbitrage condition. To see this, suppose that $\rho > \beta/\mu$. In this case, investors have no incentive to accumulate money in the morning because in the event they need money, they expect to be able to liquidate securities for a higher rate of return. Collectively, this means that the demand for real money balances goes to zero, which cannot be a part of any monetary equilibrium.[9] Suppose, alternatively, that $\rho < \beta/\mu$. In this case, investors would prefer to accumulate only money in the morning because in the event they want securities (that is, in the event they turn out to be patient), they expect to purchase these securities cheaply. However, because securities would in this case be absent, their price would be infinite–a contradiction. Thus, at the individual level investors are indifferent between holding money or capital goods in their wealth portfolios because the short-run rate of return on money and capital goods is equated, in equilibrium, through what is effectively a no-arbitrage condition.

Next, combine $\rho = \beta/\mu$ with (24) to form,

$$\pi(\beta/\mu)u'(c_1) + (1-\pi)\,Ru'(c_2) = 1 \tag{37}$$

We can also establish the following result.

Lemma 1 $z_1^+ = 0$.

Lemma 1 asserts that under [A1] and [A2], impatient investors carry no cash across periods. Lemma 1 and (34) imply,

$$c_1 = (\beta/\mu)(x+k) \tag{38}$$

There are two cases to consider depending on whether $z_2^+ \geq 0$ binds or not. Assume for the moment that $z_2^+ = 0$ binds (we will later describe a region in the parameter space where this assumption is valid). Using $\rho = \beta/\mu$ with (33) implies $x = [\pi/(1-\pi)]\,k$ which, when combined with (38) implies $k = (\mu/\beta)(1-\pi)c_1$. Next, combine this latter expression with (32) to form

$$c_2 = R(\mu/\beta)c_1 \tag{39}$$

[9]Of course, $x = 0$ is consistent with a non-monetary equilibrium, which exists in every monetary model.

Conditions (37) and (39) characterize the equilibrium allocation (c_1, c_2) when investors choose to carry no cash across periods. This case corresponds closely to the static model in Allen and Gale (2007, § 3.2) except that in our case, zero saving is derived as a property of the equilibrium in certain circumstances, rather than as something assumed to hold at all times. The relevance of this latter observation will be demonstrated shortly.

Consider next the case for which $z_2^+ \geq 0$ is slack. Condition (28) together with $\rho = \beta/\mu$ implies that c_2 is determined by,

$$Ru'(c_2) = \beta/\mu \tag{40}$$

Combine (40) with (37) to form,

$$\pi u'(c_1) + 1 - \pi = \mu/\beta \tag{41}$$

In this case, the equilibrium allocation (c_1, c_2) is determined by (40) and (41), with z_2^+ determined by (33) which, after some rearrangement, can be expressed as

$$\beta z_2^+ = c_1 - \left(\frac{\beta/\mu}{R}\right) c_2 > 0 \tag{42}$$

Here, we have what we think is an interesting result.

Proposition 2 *Given $\eta = 1$, the Friedman rule ($\mu \searrow \beta$) implements the first-best allocation (c_1^*, c_2^*) as a securities market equilibrium. First-best implementation requires that patient investors carry strictly positive amounts of money across periods.*

The proof follows immediately by evaluating (40) and (41) when $\mu = \beta$. Moreover, it is evidently critical for the result to hold that patient investors carry strictly positive money balances over time. To see this clearly, consider the first-best allocation for preferences $u'(c) = c^{-\sigma}$, $\sigma > 1$. From (7) we have $c_1^* = 1$ and $c_2^* = R^{1/\sigma}$. Evaluating (42) at the first-best allocation implies $\beta z_2^{*+} = 1 - R^{1/\sigma - 1} > 0$. Thus, the ability to save cash across periods has an important implication for the desirability of market implementation and serves to distinguish our results substantially from Allen and Gale (2007, § 3.2), which we elaborate on below.

We now characterize the properties of the securities market equilibrium more sharply. While it is possible to proceed in a slightly more general manner, in what follows we restrict ourselves to a class of utility functions that permit a closed-form solution. In particular, assume that $u'(c) = c^{-\sigma}$, where $\sigma > 1$. The tractability afforded us by this assumption will permit us to develop results in the most transparent manner possible.

Let us return to the case $z_2^+ = 0$. Since the cash constraint binds in this case, from condition (28) we know that $Rc_2^{-\sigma} > \rho$. Using (39), this latter condition can be expressed as $(R/\rho)^{1-\sigma} c_1^{-\sigma} > 1$. Now use (37) and (39) to solve for $c_1^{-\sigma} = \left[\pi\rho + (1-\pi)R^{1-\sigma}\rho^{\sigma}\right]^{-1}$. Thus, the necessary condition is given by,

$$(R/\rho)^{1-\sigma} > \pi\rho + (1-\pi)R^{1-\sigma}\rho^{\sigma} \tag{43}$$

Condition (43) will of course hold for a wide range of parameters. But there are also regions in the parameter space where it does not hold. To illustrate this possibility clearly, consider the special case $\sigma = 2$. In this case, (43) implies that the inflation rate μ needs to exceed a critical value $\mu_0 > \beta$, i.e., $\mu > \mu_0 \equiv \beta(1-\pi)/(1-\pi R) > \beta$. Thus, patient investors dispose of all their cash holdings in the afternoon securities market only if the inflation rate is sufficiently high ($\mu > \mu_0 > \beta$). Otherwise, they are willing to carry cash over time even if monetary policy operates away from the Friedman rule, at least, as long as inflation is sufficiently low ($\mu_0 > \mu > \beta$).

Let (c_1^D, c_2^D) denote the equilibrium allocation assuming [A2] and our parametric restriction. From the results above, we have

$$c_1^D = \begin{cases} (\pi\rho)^{1/\sigma}[1-(1-\pi)\rho]^{-1/\sigma} & \text{if} \quad \mu < \mu_0 \\ \left[\pi\rho + (1-\pi)R^{1-\sigma}\rho^{\sigma}\right]^{1/\sigma} & \text{if} \quad \mu \geq \mu_0 \end{cases} \tag{44}$$

$$c_2^D = \begin{cases} (R/\rho)^{1/\sigma} & \text{if} \quad \mu < \mu_0 \\ \left[\pi R^{\sigma}\rho^{1-\sigma} + (1-\pi)R\right]^{1/\sigma} & \text{if} \quad \mu \geq \mu_0 \end{cases} \tag{45}$$

where, of course, $\rho = \beta/\mu$.

Proposition 3 *Assume $\eta = 1$ and $u'(c) = c^{-\sigma}$, where $\sigma > 1$. In a securities market equilibrium, the consumption allocation possesses the following properties: (i) $c_1^D(\mu) < c_1^*$ and is strictly decreasing in μ, with $c_1^D(\beta) = c_1^*$; (ii) $c_2^D(\mu) > c_2^*$ and is strictly increasing in μ, with $c_2^D(\beta) = c_2^*$.*

Proposition 3 implies that inflation harms risk-sharing in a securities market equilibrium. The intuition for this result is as follows. First, recall the intuition developed above for the no-arbitrage condition $\rho = \beta/\mu$. Consider an exogenous increase in μ with no change in ρ so that $\rho > \beta/\mu$. The effect of this change is to induce a portfolio substitution in the morning away from cash into capital investment, an effect reminiscent of Tobin (1965). That is, capital goods can now be sold at a relatively higher price ρ relative to money, which now has a diminished rate of return β/μ. The expansion in capital investment implies a greater supply of evening output, which benefits patient investors. The lower real rate of return on money balances translates into less purchasing power for impatient investors.

It is of some interest to compare our results with Allen and Gale (2007, § 3.2) who report that the market mechanism studied here implements the first-best only in the special case of logarithmic utility. Our Propositions 1 and 2, in contrast, assert that a market mechanism implements the first-best allocation for a broader class of preferences, so long as monetary policy is set optimally.

The discrepancy in these results can be traced to the fact that our model is explicitly dynamic. Allen and Gale (2007, § 3.2), in contrast, use a version of the original Diamond and Dybvig (1983) model, which is static in nature. In our model, when the return on the short asset is equal to its social optimal level, patient investors want to save across periods. In Allen and Gale (2007, § 3.2), they are not permitted to do so.[10] The equivalent restriction

[10]Note also that in Allen and Gale (2007, § 3.2), as in Diamond and Dybvig (1983), the short-rate return is given technologically.

in our model would be to require $z_1^+ = z_2^+ = 0$ always. In this case, the allocation described in (44) and (45) for $\mu \geq \mu_0$ would now hold for the entire range of $\mu \geq \beta$, in particular,

$$\begin{aligned} c_1^D &= \left[\pi\rho + (1-\pi)R^{1-\sigma}\rho^{\sigma}\right]^{1/\sigma} \\ c_2^D &= \left[\pi R^{\sigma}\rho^{1-\sigma} + (1-\pi)R\right]^{1/\sigma} \end{aligned}$$

At the Friedman rule, $\rho = 1$, so that

$$\begin{aligned} c_1^D &= \left[\pi + (1-\pi)R^{1-\sigma}\right]^{1/\sigma} \\ c_2^D &= \left[\pi R^{\sigma} + (1-\pi)R\right]^{1/\sigma} \end{aligned}$$

In this case, $(c_1^D, c_2^D) = (c_1^*, c_2^*) = (1, R)$ if and only if $\sigma = 1$. However, as our analysis makes clear, this result seems to have less to do with the nature of preferences and more to do with an implicit restriction on saving behavior.

Proposition 3 is important because it implicitly contains the conditions necessary to obviate a role for banking in the Diamond and Dybvig (1983) model. The first condition is that monetary policy corresponds to the Friedman rule. The second condition is the availability of a freely-available competitive securities market that permits the selling of securities. If either or both of these conditions do not hold, then a bank-like institution that improves risk-sharing is likely to emerge.

Proposition 3 holds for a wider class of preferences. As shown in Lemma 1, under [A1], impatient investors do not save cash while patient investors may or may not save. The saving behavior of patient investors depends on the inflation rate: for low enough inflation rates ($\mu < \mu_0$) they save and for high enough inflation rates ($\mu \geq \mu_0$) they do not. The specific example $u'(c) = c^{-2}$ was used to derive an explicit cutoff value for the inflation rate μ_0. With the more general assumption [A1] the cutoff value may not exist; i.e., it may be the case that $\mu_0 \to \infty$, implying that patient investors always save even at very high rates of inflation.

3.2 Limited participation and banking

Proposition 2 asserts that with full market participation ($\eta = 1$), the Friedman rule ($\mu \searrow \beta$) implements the first-best allocation as a competitive monetary equilibrium. In this case, a competitive securities market constitutes a socially-efficient liquidity insurance mechanism. Proposition 3 asserts that this is no longer true for monetary policies that depart from the Friedman rule ($\mu > \beta$). In fact, we can demonstrate the following result.

Lemma 2 *The first-best allocation cannot be implemented as a securities market equilibrium for any $0 < \eta < 1$, even at the Friedman rule ($\mu \searrow \beta$).*

Intuitively, the first-best allocation requires that patient investors should receive the same consumption allocation, regardless of market access. Such an outcome is impossible, even at the Friedman rule, if a subset of investors are precluded from transforming their low-return cash balances into higher return securities (the Friedman rule does not eliminate this wedge in relative returns). It follows that if (i) monetary policy operates away from the Friedman rule and/or (ii) securities market participation is limited, then a Diamond and Dybvig (1983) style bank that replaces the securities market may provide a superior risk-sharing arrangement. We now investigate this possibility for the case $\eta = 0$, the assumption implicit in Diamond and Dybvig (1983).

In what follows, we assume that workers prefer to use cash over bank deposits.[11] We might instead have assumed that workers accept bank deposit liabilities as in Skeie (2008). In this case, the demand for government money would be driven out of circulation if it did not offer the same rate of return as private money. While we assume that workers only accept government money, our analysis would survive if we had assumed that workers prefer to hold *some* of their liquidity in the form of government money. At the end of the day, we want there to be a demand for government money since empirically we observe banks creating deposit liabilities made redeemable in government money.

A bank in this context is an intermediary that offers investors a contract (d, c_1, c_2), where d denotes an investor's initial (morning) deposit and (c_1, c_2) denotes a set of history-dependent withdrawal limits for the afternoon and evening, respectively. Agents participating in a banking arrangement are henceforth labeled *depositors*. Because depositor type is private information, the liabilities (c_1, c_2) issued by the bank will have to be made *demandable* (the early withdrawal option must be made exercisable at depositor discretion). Moreover, because depositors wanting to withdraw early will need cash to pay workers, the bank must hold cash reserves for the afternoon trading period.

The morning deposit d can be interpreted here as collateral in the form of claims issued by investors against morning output. To acquire cash, the bank sells $x \leq d$ units of this collateral to workers in exchange for $m_t^a = p_t^m x$ dollars of cash, destined to be carried as reserves by the bank into the afternoon (workers then redeem these securities for morning consumption).[12] The remaining claims to morning output $k = d - x$ are used to purchase income-generating capital goods (produced by depositors, who are in a position to manufacture such goods).

Thus, a bank effectively takes deposits d in the morning and effectively divides these resources between cash x and physical capital k, i.e., $d = x + k$. The deposit liabilities issued against these assets are assumed to be perfectly enforceable. Deposit liabilities not redeemed in the afternoon constitute *pro rata* claims against a bank's remaining assets (any residual cash and the income generated from capital goods). Because capital depreciates fully at the end of each period and because preferences are quasilinear, without loss of generality, we can restrict attention to "static" contracts.[13] We also assume that banks

[11] One way to model this choice as an equilibrium outcome is along the lines of Andolfatto (2018), where heterogeneous workers face a fixed cost of accessing the banking system. While modeling this choice here would be interesting, it is not necessary for the main results that follow and so, in the interest of simplicity, we simply assume that workers use cash.

[12] One might alternatively imagine that investors directly produce and sell consumer goods in exchange for cash, which they then deposit with their bank as part of their collective agreement.

[13] That is, we may assume without loss in generality that cash net of afternoon redemptions taken into the evening are fully paid out, along with any returns to capital. See also Berentsen, Camera and Waller

behave symmetrically in equilibrium. From this point on any reference to a *bank* should be understood to mean a representative price-taking bank unless otherwise noted.

Measured in units of afternoon output, afternoon cash reserves are given by $(\beta/\mu)\,x$ and capital goods are given by k. Let (k_2, y_2) denote the amount of capital goods and cash, respectively, used to finance c_2. Let (z_1^+, z_2^+) denote the individual purchasing power of cash carried into the next morning by impatient and patient investors, respectively. Then, the bank faces the following afternoon budget constraint

$$\pi c_1 = (\beta/\mu)\,x + r(k - k_2) - y_2 - \pi\beta z_1^+ - (1-\pi)\beta z_2^+ \tag{46}$$

The evening budget constraint is given by

$$(1-\pi)c_2 = Rk_2 + y_2 \tag{47}$$

Given assumptions [A1] and [A2], it is straightforward to demonstrate that optimality will entail the following: (i) $k_2 = k$ (capital goods are not used prematurely), (ii) $y_2 = 0$ (cash will never be used to finance evening consumption), and (iii) $z_1^+ = 0$ (impatient depositors spend all their cash in the afternoon). In what follows, we impose these restrictions beforehand to enhance the clarity of the exposition.

As an aside, note that we can think of banks here either as depositor-cooperatives, or as monopoly banks interested in maximizing their own profit. In this latter case, we could assume that banks have linear preferences over morning output. As such, we could model banks as designing deposit contracts (d, c_1, c_2) that maximize the expected welfare of depositors, charging depositors a lump-sum participation fee in the morning for services rendered. A monopoly bank would be able to charge a fee that extracted the entire investor surplus (their gains from entering into the banking arrangement net of the payoff associated with their next best alternative). Because utility is transferable here (i.e., preferences are quasilinear), the amount of surplus extracted by a monopoly bank would in no way affect the equilibrium allocation (c_1, c_2). If markets are contestable (Baumol, 1982), then the fee is, in equilibrium, bid down to the cost of banking (normalized to zero here). The point of mentioning this here is that in what follows, parameter changes that affect depositor welfare could also be interpreted as affecting bank profits to the extent that banks possess some market power.

Let $W(z)$ denote the welfare of a depositor who enters the morning with real money balances z. These balances allow a depositor to save on effort in the morning; hence, $W(z)$ is linear in z and so, $W(z) = W(0) + z$. Competition will drive banks to create a depositor base sufficiently large to diversify away idiosyncratic liquidity risk. A competitive and fully diversified bank maximizes the representative depositor's welfare by solving the following problem,

$$W(z) \equiv \max_{x,k,z_2^+} z - x - k + \pi u(c_1) + (1-\pi)[u(c_2) + \beta z_2^+] + \beta W(0)$$

subject to $z_2^+ \geq 0$, $\pi c_1 = (\beta/\mu)x - (1-\pi)\beta z_2^+$, $(1-\pi)c_2 = Rk$, and the incentive-compatibility condition

$$u(c_2) + \beta z_2^+ \geq u(c_1) \tag{48}$$

(2007).

Condition (48) ensures truthful revelation on the part of patient investors, i.e., they will prefer to wait until the evening to withdraw their deposits.[14]

Let $(1-\pi)\beta\psi$ denote the Lagrange multiplier associated with the non-negativity constraint $z_2^+ \geq 0$, and ϕ the Lagrange multiplier for the incentive constraint (48). The conditions for optimality are given by

$$\begin{aligned}
(1-\phi/\pi)\,(\beta/\mu)u'(c_1) &= 1 && (49)\\
[1+\phi/(1-\pi)]\,Ru'(c_2) &= 1 && (50)\\
1+\phi\left[1/(1-\pi)+u'(c_1)/\pi\right]+\psi &= u'(c_1) && (51)
\end{aligned}$$

Let us conjecture (and then verify) that the Lagrange multiplier $\phi = 0$. In this case, the banking allocation (c_1^B, c_2^B) is determined by

$$\begin{aligned}
(\beta/\mu)u'(c_1^B) &= 1 && (52)\\
Ru'(c_2^B) &= 1 && (53)
\end{aligned}$$

From (52)–(53), $c_2^B > c_1^B$, since $R > \beta/\mu$. Thus, since $z_2^+ \geq 0$, it follows that (48) is slack, which confirms our conjecture that $\phi = 0$.

It is immediately evident from (52)–(53) that the banking equilibrium implements the first-best allocation at the Friedman rule. Moreover, because of the quasilinearity of preferences, the evening allocation is invariant to inflation. The afternoon allocation is strictly decreasing in the rate of inflation, reflecting the usual inflation-tax effect on cash goods. For $u'(c) = c^{-\sigma}$, $\sigma > 1$, conditions (52)–(53) imply that the bank allocation is given by,

$$\begin{aligned}
c_1^B &= (\beta/\mu)^{1/\sigma} \leq c_1^* && (54)\\
c_2^B &= R^{1/\sigma} = c_2^* && (55)
\end{aligned}$$

Proposition 4 *Assume $u'(c) = c^{-\sigma}$, $\sigma > 1$. Then the equilibrium allocation in the banking economy possesses the following properties: (i) $c_1^B(\mu) < c_1^*$ and is strictly decreasing in μ, with $c_1^B(\beta) = c_1^*$; (ii) $c_2^B = c_2^*$ and is invariant to μ.*

Finally, it is of some interest to note that under a banking arrangement, it is never optimal for either banks or depositors to carry "excess" cash over time. Combining (52) with (51) implies

$$\psi = \mu/\beta - 1 \qquad (56)$$

Thus, for inflation rates satisfying $\mu > \beta$, we have $z_2^+ = 0$. When $\mu = \beta$ we have $z_2^+ \geq 0$, but we can set $z_2^+ = 0$ without loss of generality. This is in contrast with the securities market equilibrium where patient investors found it optimal to carry cash balances over time for inflation rates $\mu \leq \mu_0$.

[14] A patient depositor prefers to wait until the evening if the value of doing so is at least as large as the value of misrepresenting himself, i.e., declaring to be impatient and withdrawing in the afternoon. The bank contract needs to provide incentives for truthful revelation of type. Formally, this implies the incentive compatibility condition, $u(c_2) + \beta W(z_2^+/(1-\pi)) \geq u(c_1) + \beta W(0)$, which given $W(z) = W(0) + z$ simplifies to (48).

The allocation under banking can be viewed as *constrained-efficient* in the sense that the banking equilibrium replicates what a planner facing a rate of return of β/μ between morning and afternoon would implement. This is analogous to assuming the rate of return on the short-term technology in a standard Diamond-Dybvig setup is less than one. In that case banks implement the first-best for any given rate of return on the short-term technology. Of course, in our model the short-term rate of return is determined by government policy rather than technology, so that an efficient monetary policy renders banks redundant.

3.3 Banks vs markets

In this section, we compare equilibrium allocations under a securities market with full participation (44)–(45) with that of a banking system (54)–(55) assuming that the two systems do not interact in any way. The first thing to note is that when $\eta = 1$, both institutions deliver the optimal risk-sharing arrangement when monetary policy follows the Friedman rule. Away from the Friedman rule, the banking arrangement offers superior risk-sharing than the securities market and hence, higher *ex ante* welfare for depositors.

Proposition 5 *Investor welfare in a banking equilibrium is greater than in a securities market equilibrium with $\eta = 1$; that is, $W(z) > V(z)$, for all $z \geq 0$.*

The difference in bank and market allocations depends on whether inflation is below or above the threshold μ_0 we identified above. For impatient investors we have

$$c_1^B/c_1^D = \begin{cases} \pi^{-1/\sigma}[1-(1-\pi)(\beta/\mu)]^{1/\sigma} & \text{if} \quad \mu < \mu_0 \\ \left[\pi + (1-\pi)(R\mu/\beta)^{1-\sigma}\right]^{-1/\sigma} & \text{if} \quad \mu \geq \mu_0 \end{cases}$$

Given [A2], it is easy to show that $c_1^B > c_1^D$ for all $\mu > \beta$. Also, from the expressions above, it is clear that the distance between c_1^B and c_1^D increases with inflation.

Since evening consumption is at its first-best level in the banking equilibrium, the allocation here dominates the market allocation for all inflation rates. In particular, $c_2^B = c_2^* < c_2^D$ for all $\mu > \beta$. Furthermore, given Proposition 3 the distance between c_2^B and c_2^D also increases with inflation. That is, as inflation rises, the distance between the consumption of impatient and patient investors grows larger in the market economy than with banking, so that risk-sharing becomes relatively poorer. Thus, even though welfare under both types of arrangements suffers with higher inflation rates, banking becomes relatively more valuable as inflation increases.

Suppose that monetary policy operates away from the Friedman rule. Then Proposition 5 asserts that a banking system dominates a financial market–*even when all investors have costless access to the financial market*. What is it that banks can do that markets cannot? The answer is that banks are able to pool idiosyncratic liquidity risk and that pooling cash reserves in this manner has benefits in a high inflation environment. Recall that for efficiency to be achieved in a financial market, investors had to be willing to carry cash across time. They are less willing to do so when inflation is too high. Their collective attempt to economize on real cash balances in this case further distorts investor portfolio decisions.

4 Coexistence of banks and financial markets

In the analysis above, we studied two monetary economies, one with a securities market and one with a banking system. According to Proposition 5, banks are generically strictly better than markets at delivering liquidity insurance. Proposition 5, however, was derived under the assumption that bank depositors cannot access a securities market after types are revealed. While this assumption is commonly employed in the literature beginning with Diamond and Dybvig (1983), it is not innocuous. Jacklin (1987), Haubrich (1988), von Thadden (1997) find that the availability of an *ex post* securities market typically destroys the *ex ante* insurance gains made available by banking. All of these papers, however, assume full participation. The limited participation assumption employed by Diamond (1997) has the effect of tempering this dramatic result–banking arrangements may continue to outperform what is achievable with securities markets.

Thought experiments in the spirit of Jacklin (1987) do not actually model coexistence as much as they consider the *threat* of coexistence. That is, they typically consider how the incentives for a mass defection of impatient and patient bank depositors trading competitively among themselves constrain the structure of the bank contract *ex ante*. The constrained-efficient banking arrangement discourages depositors (leaving them at least indifferent) from utilizing what turns out to be, in equilibrium, a non-existent market. In what follows we take a different route.

4.1 Banking in the presence of financial markets

Consider a depositor with contract (c_1^B, c_2^B) solving (52)–(53), who turns out to be patient in the afternoon. Since the incentive-compatibility condition (48) holds, he clearly prefers to withdraw his funds in the evening. However, if the depositor has access to a securities market then, instead of spending or saving the cash he withdraws in the afternoon, he could use it to purchase securities at price ρ. The question is how the opportunity for securities market trading impinges on the ability of banks to offer liquidity insurance.

Patient depositors with market access can withdraw c_1^B units of (real) money in the afternoon, exchange it for capital at price ρ and earn a return R in the evening. Assuming they spend all their money, the resulting evening consumption is equal to $(R/\rho)c_1^B$. If this amount is larger than c_2^B, patient depositors will prefer to withdraw their funds in the afternoon rather than wait until the evening. Thus, patient depositors with market access have an incentive to withdraw early when the price of capital is low enough. We show this result formally in the statement below.

Proposition 6 *Consider the Diamond-Dybvig bank contract* (c_1^B, c_2^B) *that solves* (52)–(53). *Consider any securities market with price* $0 < \rho \leq \beta/\mu$. *Then a patient depositor with market access would prefer to withdraw his deposit in the afternoon.*

The thought experiment we have in mind is as follows. Consider the securities market equilibrium described above for a given pair (μ, η). Associated with this equilibrium is a securities price, which by Proposition 1 satisfies $0 < \rho < \beta/\mu$ when $\eta < 1$ and $\rho = \beta/\mu$

when $\eta = 1$. Next, consider a subset of this population–formally, a countable infinity of investors drawn randomly from the general population. If this small subset of individuals can interact with the rest of the population–which we assume they do–then they can properly be considered a small open economy. Assume that this subset has the ability to form a bank (the remainder of the population remains unbanked). The question we ask is how the availability of a securities market with exogenous price ρ influences the structure of the bank contract for this subset of investors.

Imagine that associated with the technology of bank formation is a separate parameter $\tilde{\eta} \in [0,1]$ that governs the ease of market access by patient depositors. It could be that $\tilde{\eta} = \eta$, but in general this need not be the case. For example, $\tilde{\eta} = 0$ could be thought of as a bank with the power to render its relationship with its depositors fully exclusive. Alternatively, $\tilde{\eta} = 1$ could be thought of as a bank accessing the securities market on behalf of its depositors, say, in the manner of a money market fund.

The availability of an active securities market complicates matters for the bank in an interesting way. For the case $\tilde{\eta} = 0$, the bank would be able to implement the Diamond-Dybvig solution (c_1^B, c_2^B) described in (52)–(53). However, for any $\tilde{\eta} > 0$, Proposition 6 asserts that patient depositors with market access would want to spend their money on cheap securities, rendering the Diamond-Dybvig solution non-implementable. If this is the case, then what is the solution to the bank's optimal risk-sharing arrangement when $\tilde{\eta} > 0$ and $0 < \rho \leq \beta/\mu$? In the analysis that follows, we assume:

[A3] $0 < \tilde{\eta} \leq 1$ and $0 < \rho \leq \beta/\mu$

While the possibility of *ex post* trade may reduce *ex ante* risk-sharing, it also provides a profitable investment opportunity since capital can now be acquired cheaply in the afternoon securities market. The question is then, whether the bank should discourage its depositors from withdrawing funds early to exploit this profitable trading opportunity. Discouraging early redemptions by patient depositors with market access is always feasible. In particular it would require reducing c_1 (effectively, the short term deposit rate) and increasing c_2 (the long term deposit rate) to a point that renders early redemption unattractive. Such an action, however, obviously comes at the cost of less risk-sharing for depositors. As it turns out, one can demonstrate that under [A3], an efficient bank contract will not discourage patient depositors with market access from withdrawing funds early. This result is shown formally in Appendix C.

The constrained-efficient bank contract is now characterized. The bank sets aside m^a dollars in reserve in the morning or, measured in units of afternoon consumption, $c_1 = (m^a/p^a)/\left[(\pi + (1-\pi)\,\tilde{\eta})\right]$. In particular, note that the bank must in this case accumulate cash reserves not only for impatient depositors but also for the measure of patient depositors $(1-\pi)\,\tilde{\eta}$ it expects to have market access.

Let $\tilde{k}$ denote the amount of securities purchased by a patient depositor with market access. Since the nominal value of purchased capital cannot exceed the cash withdrawn from the bank, we have $q\tilde{k} \leq m^a/\left[(\pi + (1-\pi)\,\tilde{\eta})\right]$ or, in real terms, $\rho\tilde{k} \leq c_1$ (recall that $\rho \equiv q/p^a$). Any unspent cash is carried into the next morning; hence, $\beta\tilde{z}_2^+ = c_1 - \rho\tilde{k} \geq 0$.

The capital goods purchased in the securities market generate $\tilde{c}_2 = R\tilde{k}$ consumption goods in the evening. We can combine these latter two expressions to write the budget constraint of the patient agents with market access as follows:

$$\beta \tilde{z}_2^+ = c_1 - (\rho/R)\tilde{c}_2 \geq 0 \tag{57}$$

As before, the bank faces an afternoon budget constraint and an evening budget constraint. The evening budget constraint is $(1-\pi)(1-\tilde{\eta})c_2 = Rk$. The afternoon budget constraint is derived as follows. At the beginning of the afternoon, the bank holds m^a units of money and each depositor with market access receives $m^a/(\pi + (1-\pi)\tilde{\eta})$ dollars. Hence, in real terms the budget constraint satisfies $(\beta/\mu)x = [\pi + (1-\pi)\tilde{\eta}]c_1$.

Let $W(z)$ denote the value of beginning the morning with real money balances z. These balances allow a depositor to save on effort the following morning; hence, z enters linearly and we again obtain $W(z) = W(0) + z$. Thus, we can use (57) to write $\beta W(\tilde{z}_2^+) = \beta W(0) + c_1 - (\rho/R)\tilde{c}_2$. The choice problem facing a bank on behalf of a depositor who enters the morning with real money balances z can therefore be expressed as

$$W(z) \equiv \max_{x,k,\tilde{c}_2} z - x - k + \pi u(c_1) + (1-\pi)\tilde{\eta}[u(\tilde{c}_2) + c_1 - (\rho/R)\tilde{c}_2] + (1-\pi)(1-\tilde{\eta})u(c_2) + \beta W(0) \tag{58}$$

subject to the bank's budget constraints $(\beta/\mu)x = [\pi + (1-\pi)\tilde{\eta}]c_1$ and $(1-\pi)(1-\tilde{\eta})c_2 = Rk$, the budget constraint of a patient depositor with market access (57) and the incentive-compatibility conditions

$$u(\tilde{c}_2) + c_1 - (\rho/R)\tilde{c}_2 \geq u(c_2) \geq u(c_1) \tag{59}$$

Note that if (57) binds, then $\beta \tilde{z}_2^+ = c_1 - (\rho/R)\tilde{c}_2 = 0$ and the incentive compatibility conditions in (59) simplify to $\tilde{c}_2 \geq c_2 \geq c_1$.

Let $(1-\pi)\tilde{\eta}\tilde{\zeta}$ denote the Lagrange multiplier associated with the inequality constraint (57), $c_1 - (\rho/R)\tilde{c}_2 \geq 0$. If the incentive-compatibility conditions in (59) are slack, the first-order conditions imply

$$\pi u'(c_1) + (1-\pi)\tilde{\eta}(R/\rho)u'(\tilde{c}_2) = (\mu/\beta)[\pi + (1-\pi)\tilde{\eta}] \tag{60}$$

$$Ru'(c_2) = 1 \tag{61}$$

$$(R/\rho)u'(\tilde{c}_2) = 1 + \tilde{\zeta} \tag{62}$$

It is of some interest to compare (60)–(62) to the conditions characterizing the optimal bank contract absent financial markets (52)–(53). Notice that condition (61) corresponds to (53), so that the consumption payoff received by the patient depositor without market access remains unaffected. Thus, the new dimension to the problem now involves structuring the asset portfolio between cash and capital in a manner that trades off the returns available for impatient depositors and the set of patient depositors who gain market access. Note too that the marginal cost of accumulating cash reserves is higher in this case, given that cash is needed to satisfy both impatient depositors and patient depositors with market access. We now verify that the incentive compatibility constraints are satisfied and further characterize the constrained-efficient deposit contract.

Proposition 7 *The deposit contract $(c_1, c_2, \tilde{c}_2)$ characterized by (57)–(62) is incentive-compatible, that is, $u(\tilde{c}_2) + c_1 - (\rho/R)\tilde{c}_2 \geq u(c_2) \geq u(c_1)$. Furthermore, $c_1 < c_1^B < c_1^* < c_2 = c_2^B = c_2^* < \tilde{c}_2$, where (c_1^B, c_2^B) solves (52)–(53).*

How does the bank deposit contract change in the presence of a securities market? Patient depositors without market access receive the same evening allocation ($c_2 = c_2^B = c_2^*$). Patient depositors with market access profit from the low market price of capital ($\rho \leq \beta/\mu$). They achieve a higher level of consumption ($\tilde{c}_2 > c_2^B = c_2^*$) by withdrawing their money and using it to buy cheap securities. The higher consumption for patient depositors with market-access comes at the cost of lower consumption for impatient depositors ($c_1 < c_1^B < c_1^*$).

4.2 Depositor welfare

Recall that for the Diamond-Dybvig bank contract that solves (52)–(53), welfare is strictly decreasing in the rate of inflation, μ. It is interesting to note that the effect of inflation on depositor welfare when depositors have access to the securities market is generally ambiguous. A higher rate of inflation is likely to reduce the price of securities (this is obviously true for the case $\rho = \beta/\mu$). Using the envelope theorem on the welfare function (58) we see that welfare is decreasing in μ and ρ. Consequently, if increasing μ leads to lower ρ, welfare can potentially rise. That is, while inflation leads to reduced risk-sharing, it also depresses securities prices. The opportunity for bank depositors to pick up cheap securities (high rate of return R/ρ) is clearly welfare-improving.

Proposition 8 *If $(d\rho/d\mu)(\mu/\rho) \geq -1$ then welfare of depositors is strictly decreasing in inflation.*

The result above reports a sufficient condition under which inflation affects depositors adversely. This condition, expressed as the elasticity of the securities price with respect to inflation, is trivially satisfied when ρ is nondecreasing in μ; when ρ is decreasing in μ, the condition places a restriction on how fast the securities price falls with inflation.[15] In all these cases, the best policy is to implement the Friedman rule, though this is typically not enough to achieve the first-best allocation.

The availability of a securities market affects how banks provide liquidity insurance for depositors. When $\eta = 1$ and hence, $\rho = \beta/\mu$, we find that the welfare of depositors is strictly decreasing in $\tilde{\eta}$. The outcome is intuitive: higher values of $\tilde{\eta}$ further impinge on a bank's ability to provide liquidity insurance, as more and more depositors withdraw in the afternoon and access the securities market. When $0 < \eta < 1$, and so $0 < \rho < \beta/\mu$, we find that $\tilde{\eta}$ has an ambiguous effect on depositor welfare. The negative effect we just described for the case $\eta = 1$ is still present. However, now patient depositors with market access benefit from a low securities price. Lower values of ρ (say, due to lower η) imply a higher afternoon-to-evening rate of return for patient depositors accessing the market,

[15] When $\rho = \beta/\mu$ (as in Section 3) it is easy to show that $(d\rho/d\mu)(\mu/\rho) = -1$, which satisfies the condition in Proposition 8.

R/ρ. We find that the effect of $\tilde{\eta}$ on welfare can be strictly decreasing (for sufficiently high values of ρ), non-monotone (first decreasing and then increasing) or strictly increasing (for sufficiently low values of ρ).[16]

When $0 < \rho < \beta/\mu$ it is possible for depositors with $0 < \tilde{\eta} \leq 1$ to obtain higher *ex ante* welfare than with the Diamond-Dybvig bank ($\tilde{\eta} = 0$). It may even be possible to obtain higher welfare than the first-best allocation. Note, however, that this welfare gain is achieved at the expense of welfare for unbanked investors, for which a low securities price adversely affects the amount of insurance provided by markets.

A special case to consider is when $\tilde{\eta} = 1$. Since now all depositors withdraw their deposits in the afternoon, the bank only holds cash in its portfolio, essentially becoming a narrow bank. What we can say here depends on the exact interpretation of $\tilde{\eta}$. If $\tilde{\eta}$ is interpreted as representing something intrinsic to the depositor, then for the case of $\tilde{\eta} = 1$, the bank becomes redundant since it cannot improve on what investors might otherwise achieve using the securities market on their own by following the same cash-only investing strategy. If $\tilde{\eta}$ is interpreted as something that stems from the depositor-bank relationship, then $\tilde{\eta} = 1 \geq \eta$ implies that the bank acts as a money market mutual fund and can achieve a superior allocation relative to what investors could achieve on their own (due to the bank's higher market-participation rate).

4.3 Optimal monetary policy

As described above, there may be a tension between the welfare of unbanked investors and depositors. The next result shows that when banks and securities markets coexist, the first-best allocation can be implemented in equilibrium if and only if monetary policy is at the Friedman rule *and* (unbanked) investors have frictionless access to the market.

Proposition 9 *Consider a competitive banking sector offering deposit contracts characterized by* (57)–(62) *and a securities market with* $\eta \in (0, 1]$. *The first-best allocation is implementable (for banked depositors and unbanked investors) if and only if* $\eta = 1$ *(unbanked investors have full access to securities markets) and* $\mu \searrow \beta$ *(monetary policy is at the Friedman rule). Furthermore, first-best implementation requires that patient investors and depositors with market access carry strictly positive amounts of money across periods.*

In the proof of Proposition 9, we show that when $\mu = \beta$ and $\eta = 1$, $\rho = 1$ and the savings constraint does not bind, i.e., $\tilde{\zeta} = 0$. Then, from (60)–(62), it immediately follows that $c_1 = c_1^*$ and $c_2 = \tilde{c}_2 = c_2^*$. Given our previous analysis on market and banking arrangements, it is perhaps not surprising that the first-best allocation is attainable at the Friedman rule. However, it is important to note that the result holds only if unbanked investors have perfect market access, i.e., $\eta = 1$. In contrast, the probability of market access for depositors, $\tilde{\eta}$ plays no role in determining optimality of the Friedman rule.

Since frictionless markets and banks in isolation each implement the first-best at the Friedman rule, it may seem natural to expect the same result to apply when the two systems

[16] These results are based on numerical examples using analytical solutions to the bank problem. Calculations are available upon request.

coexist. However, at the Friedman rule, patient depositors with market access still find it optimal to withdraw their bank deposits in the afternoon when $\rho = 1$ (see Proposition 6). That is, patient depositors do not behave as they do when the securities market is not available. How can this behavior be consistent with first-best implementation? Proposition 9 again highlights the role of savings: in order to implement efficient liquidity insurance, it is critical that depositors have an opportunity to save across periods. In effect, they use the withdrawn funds to purchase only enough securities to finance c_2^* and carry the remaining cash into the next period, so that they won't have to work as hard to rebalance their portfolios.

4.4 Related literature

It is some interest to compare our results relative to Farhi, Golosov and Tsyvinski (2009) and a closely related paper by Allen and Gale (2004). These authors highlight the interaction of two key frictions: private information over preference types and unobservable side trades. Our paper adds a dynamic dimension and an additional friction–a lack of commitment over debt repayment–that generates a demand for money.

In particular, Farhi, Golosov and Tsyvinski (2009) demonstrate the usual underprovision of liquidity that occurs when *ex post* trading cannot be discouraged. They then demonstrate how an intervention in the form of a broad-based minimum reserve requirement can be selected such that a banking equilibrium with side trades implements the efficient risk-sharing allocation. The intervention works through a general equilibrium effect. In particular, a legislated increase in bank reserves (think of x in our model) has the effect of lowering the equilibrium interest rate on the long-maturity instrument (this is equivalent to lowering R/ρ or increasing ρ in our model–see equation 36). This, in turn, has the effect of discouraging patient depositors from cashing out early for the purpose of arbitrage (i.e., by re-investing the proceeds in the long-maturity instrument).

In our model, the rate of return on the long-maturity instrument is related to the inflation rate, as argued above. Thus, our preferred intervention–the Friedman rule–shares the flavor of the Farhi, Golosov and Tsyvinski (2009) result. In particular, by running a deflation, the monetary authority increases the rate of return on the short-instrument (cash) and lowers the rate of return on the long-instrument (securities). This is exactly the type of relative price distortion that is necessary to correct the externality generated by the market frictions in this environment.

Our analysis is consistent with the notion that securities markets potentially inhibit the formation of superior liquidity risk-sharing arrangements that could be offered by banks with the benefit of exclusive depositor relationships.[17] Unlike Jacklin (1987), however, our bank is not rendered redundant by the existence of a securities market because, like Diamond (1997), securities markets are subject to limited participation. Our mechanism,

[17] Jacklin (1987) and Diamond (1997) also consider the restrictions placed by the possibility of an *ex ante* deviation where an investor bypasses the bank and directly acquires only capital in the morning. If he turns out to be patient he enjoys the high return to his large capital investment. If he turns out to be impatient he uses the securities market to sell it. As it turns out, this deviation places no additional restrictions in our model as the bank is already internalizing the presence of an active securities market.

however, appears somewhat different than in Diamond (1997). In his setup, as in ours, the equilibrium securities price is too low because some buyers are shut out of the market. The emergence of a bank sector in his model has a general equilibrium effect, it evidently improves the operation of the securities market. In our setup, this general equilibrium effect is absent. Our bank can increase the welfare of its depositors through enhanced risk-sharing and by permitting its depositors the option to withdraw for purposes unrelated to consumption (in particular, to exploit a profitable financial opportunity).

We are left with the unexplored question of how general equilibrium effects may alter the conclusions derived in our model. What, if anything, would change if depositors instead constituted a measurable subset of investors? We do not formally study this scenario, but it seems easy enough to sketch out its likely properties and argue that we would not expect anything to change qualitatively.

First, note that the only depositors accessing the securities market are patient depositors that withdraw cash from in the bank in the afternoon. Thus, the amount of cash relative to securities being traded in the market would increase with the measure of these depositors accessing the market, putting upward pressure on the securities price, ρ. However, note that when $\eta = \tilde{\eta} = 1$, i.e., a full measure of patient investors (both unbanked and banked) participate in the securities market, the securities price is pinned down by the no-arbitrage condition, $\rho = \beta/\mu$. Less than full participation implies less cash is exchanged in the securities market, putting downward pressure on the price. In other words, the securities price would remain within the bounds we established above, $0 < \rho \leq \beta/\mu$, and so, [A3] would continue to hold.

Second, a representative bank and representative unbanked investor are likely to hold different portfolios (x, k). For binding cash constraints, some variation of the pricing equation (36) will hold, which asserts that the securities price is directly increasing in η and increasing in the aggregate ratio of cash-to-capital in the morning wealth portfolio. Judging by the results derived above, this latter object is likely decreasing in the inflation rate, so that the equilibrium securities price will continue to be decreasing in the rate of inflation.

5 Bank runs and securities markets

Consistent with the literature we have cited repeatedly above, the analysis here supports the idea that banks are generally superior to markets as mechanisms for delivering liquidity insurance. However, to the best of our knowledge no one has asked how the availability of markets interacts with bank sector fragility.[18] The main purpose of Diamond and Dybvig (1983) was, of course, to formalize the notion of bank sector "fragility" and the rationale for deposit insurance. How are the results we derived above sensitive to the existence of bank run equilibria?

In what follows, we assume a form of contractual incompleteness that gives rise to the

[18] An exception here is Jacklin (1987) who shows that for the special preferences assumed by Diamond and Dybvig (1983), a market mechanism implements the first-best so that banks (and their associated fragility) can be dispensed with.

existence of a *bank run*–a situation in which all patient depositors without market access misrepresent themselves as wanting cash in the afternoon. In particular, we assume that the bank contract described above does not anticipate the possibility of a run. Moreover, we assume that banks cannot commit to suspending redemptions after cash reserves are depleted. Instead, we assume that if the early redemption promise cannot be met for every depositor requesting early withdrawal, then the bank becomes bankrupt and is forced to disperse its assets in some prescribed manner. Below, we consider different ways in which bankruptcy is resolved.

An important property of the bank contract analyzed in Section 4 is that, when $\tilde{\eta} < 1$, the afternoon market value of the bank's asset portfolio is not high enough to cover all its obligations in the event that all depositors decide to withdraw in the afternoon. That is, even in the best-case scenario in which the bank is able to liquidate all its assets at market prices, it would not be able to keep its promise to depositors in the event of a bank run. In contrast, when $\tilde{\eta} = 1$, all depositors withdraw in the afternoon and thus, the bank only holds cash. In this case, the bank is run-proof as the value of its liabilities equals the value of its assets.

Lemma 3 *The deposit contract characterized by* (57)–(62) *implies: (i)* $(\beta/\mu)x + \rho k < c_1$ *when* $0 < \tilde{\eta} < 1$*; and (ii)* $(\beta/\mu)x = c_1$ *and* $k = 0$ *when* $\tilde{\eta} = 1$.

From this point on, we focus on the case $0 < \tilde{\eta} < 1$, when the market value of the bank's assets are not enough to satisfy the event of a bank run.

Note that the equilibrium securities price $0 < \rho \leq \beta/\mu$, which the bank faced when designing the deposit contract, continues to prevail in the event of a bank run (recall that our bank sector is small relative to the rest of the economy). This price ρ can be thought of as an upper bound for what might happen in general equilibrium where a collapsing bank sector has a measurable impact on the economy. To see this, note that the bank needs to liquidate capital and obtain cash in order to meet deposit redemptions. If there are any general equilibrium effects of a bank run in the securities market, they will put downward pressure on ρ, due to the simultaneous increase in the supply of securities and the demand for cash. That is, the market value of the bank's portfolio may be even lower than that assumed in Lemma 3. General equilibrium effects are not central to our results, so we ignore them in our discussion below.

The resolution of a bank run depends on whether the bank itself has access to the financial market or not. Even in the case that is does, it may not be optimal to access the market and convert claims to capital into cash. We consider each case in turn.

5.1 Asset liquidation

Suppose that banks are forced to liquidate (that is, convert into cash–not scrap) their assets in the event of bankruptcy. Given market price $0 < \rho \leq \beta/\mu$, the bank sells k units of investment for ρk units of cash (measured in units of afternoon consumption). At the end of the liquidation process, the bank has real money balances totalling $(\beta/\mu)x + \rho k$, which is then dispersed on a *pro rata* basis to depositors (all of whom are asking for cash in the afternoon). By Lemma 3, $(\beta/\mu)x + \rho k < c_1$ and so, the bank has insufficient cash to honor all its obligations.

Let us consider the best case scenario in which funds are dispersed to depositors that very afternoon.[19] What do depositors do with the cash they receive? The answer depends on what trading opportunities depositors have. Impatient depositors simply spend their cash in the afternoon goods market. The measure $1 - \tilde{\eta}$ of patient depositors that do not access the securities market are compelled to spend their cash for afternoon goods as well, which they store into the evening; hence, they obtain the same consumption as impatient depositors. The measure $\tilde{\eta}$ of patient depositors who gain access to the securities market can sell all their cash in exchange for capital goods; however, depending on circumstances, they may want to carry over some cash to the following period. Let $(c_1^l, c_2^l, \tilde{c}_2^l)$ be the allocation resulting from a liquidation of the bank's asset portfolio, where the superscript l denotes "liquidation." We next establish the properties of this allocation.

Proposition 10 *Consider a deposit contract* $(c_1, c_2, \tilde{c}_2)$*, as characterized by* (57)–(62) *when* $0 < \tilde{\eta} < 1$*. In the event of a bank run, if the bank liquidates its asset portfolio and disperses the proceeds pro rata among all depositors, then all depositors receive* $(\beta/\mu)x + \rho k$ *units of cash in the afternoon and the resulting allocation* $(c_1^l, c_2^l, \tilde{c}_2^l)$ *has the following properties:* $c_1^l = c_2^l < c_1 < c_2$*; and* $\tilde{c}_2^l = \tilde{c}_2 \leq (R/\rho)c_1^l$ *if* $(R/\rho)u'((R/\rho)c_1^l) \leq 1$*, and* $\tilde{c}_2^l < \tilde{c}_2$ *otherwise.*

By Lemma 3, all depositors are subject to a "haircut" on their deposits in the event of a bank run. Proposition 11 states that under an asset liquiditation policy, all depositors will in general also obtain lower consumption. There is one exception. It could be the case that patient depositors with market access decide to save money across periods, both prior to and subsequent to the run. In this case these depositors are willing and able to dip into their savings in the event of a run to stabilize their afternoon consumption spending, i.e., $\tilde{c}_2^l = \tilde{c}_2$. Of course, reduced savings here means greater effort the next morning (to rebalance depleted wealth).[20]

5.2 Clearinghouse certificates

In the previous section, the bank liquidated claims to capital and then dispersed its cash proceeds and reserves on a *pro rata* basis. An alternative protocol is to disperse cash and and direct claims to capital *separately* on a *pro rata* basis, leaving depositors with the option of whether to hold or liquidate claims to capital (which are now equivalent to securities). In effect, this protocol converts debt to equity, an operation sometimes used in the business sector when firms are under financial distress. In the context of banking, a similar practice was used during the U.S. National Banking Era (1863-1914); see Gorton (1988). In particular, mass redemption events were often dealt with by having banks coalesce into a

[19] In reality, the liquidation process takes time and depositors wanting their funds would be forced to wait, with any added delay adding to the welfare losses associated with bank insolvency described here.

[20] This result that cash savings bears the brunt of the shock depends on quasilinear preferences. With nonlinear preferences, $\tilde{c}_2^l$ would adjust along with cash savings.

consolidated entity, suspending cash redemptions (after significant cash had left the banking system), and issuing clearinghouse certificates representing claims against the assets of the consolidated entity. These certificates evidently circulated in secondary markets and once the crisis passed, operations returned to normal. In the same spirit, assume that the bank disperses its cash holdings, $(\beta/\mu)x$ and claims to capital k (certificates) on a *pro rata* basis. The implications of following this resolution protocol relative to the one studied earlier are summarized in Proposition 11.

Proposition 11 *Consider a deposit contract* $(c_1, c_2, \tilde{c}_2)$ *as characterized by* (57)–(62) *when* $0 < \tilde{\eta} < 1$. *In the event of a bank run, if the bank disperses its cash reserves and certificate claims to remaining assets on a pro rata basis so that all depositors receive* $(\beta/\mu)x$ *units of cash and* k *certificates, then the resulting allocation* $(c_1^h, c_2^h, \tilde{c}_2^h)$ *has the following properties:* $c_1^h = c_1^l$; $c_2 > c_2^h > c_2^l$; *and* $\tilde{c}_2^h = \tilde{c}_2^l$.

Under this protocol, impatient depositors receive the same allocation they would under the liquidation procedure. This latter result is, however, sensitive to the Diamond (1997) assumption that impatient depositors retain costless access to the securities market. As such, they are able to sell their certificates for the cash they need to purchase afternoon consumption. Likewise, the measure $\tilde{\eta}$ of patient depositors who gain access to the securities market spend their cash on certificates which they can redeem later in the evening for output, resulting in the same consumption they would have attained under asset liquidation.

The group with a different outcome consists of the patient depositors without market access. These depositors now obtain a higher consumption allocation $(\beta/\mu)x + Rk$ instead of $(\beta/\mu)x + \rho k$. Since $R > \rho$, their marketable wealth is higher. Put differently, since these depositors are patient, they prefer holding securities to term (made possible by the present arrangement) rather than having the bank liquidate the underlying claims to capital on their behalf at a discount ($\rho < R$). Thus, the clearinghouse arrangement described here strictly dominates the asset liquidation protocol described earlier, though admittedly *strict* dominance is a byproduct of the assumption that impatient depositors have full access to the securities market. If impatient depositors had less than full access, then those without market access would have to liquidate their securities at a deep discount (here, they would scrap their claims to capital) for the very low rate of return r. In this more general case, the two protocols described here would exhibit a more interesting trade-off.

5.3 Insurance vs. stability

Banks offer superior liquidity insurance relative to markets. But if banks are subject to runs, then the benefit of this added insurance should be weighed against the costs of possible disruptions in the payments system. It is interesting to note that conditional on a bank run, the cost of the subsequent disruption is smaller for economies in which depositors have easier access to securities markets. In the limiting case for which there is perfect access to securities markets, fragility is inconsequential. The development of securities markets then is predicted to help smooth the adverse consequences associated with bank runs, possibly eliminating the phenomenon altogether even without the aid of a lender-of-last-resort.

On the other hand, our analysis suggests that stability through improved access to securities markets may come at a cost. In our model, as access to securities markets becomes progressively less costly over time, the ability of banks to provide superior liquidity insurance may diminish accordingly. The quantitative impact of these developments is predicted here to depend largely on the inflation rate regime. For high inflation rate regimes, the economic benefits of liquidity insurance through the banking sector is higher. But for low inflation rate regimes, the welfare gains of banks over markets diminish (see Section 3.3). We conclude that a policy of keeping inflation low and stable as access to financial markets improves over times seems like a good way to promote financial stability while promoting efficient liquidity insurance arrangements.

The discussion above is related to the literature on the stability role of the maturity structure of government debt (e.g., Greenwood, Hanson and Stein, 2015 and Krishnamurthy and Vissing-Jorgensen, 2015). In this literature, providing more short-term safe assets crowds out production of short-term deposits and increases financial stability. In our model, there is only one government security–a zero-interest bond (money). This government security is useful in transactions.

One could imagine adding to the model a long-term illiquid security that would be priced at its "fundamental" value. Away from the Friedman rule, the short-term debt (money) would trade at a premium relative to the long-term security. That is to say, the exchange medium (or collateral asset) would be "scarce." The scarcity of liquidity would motivate the formation of fractional reserve banks. The effect of this is to expose the economy to financial instability. The effect of lowering the inflation rate is to increase the real rate of return on short-term debt (money), encouraging banks and other agencies to become satiated with liquidity.

At the Friedman rule, banks would be willing to hold "excess reserves," i.e., voluntarily hold enough cash to meet the highest possible redemption activity so that runs are no longer possible.[21] The financial system becomes stable in this sense. However, there is no meaningful "crowding out" of private production of short-term deposits–all that happens in our model is that bank deposits become more highly-backed by cash.

5.4 Lender of last resort

The Diamond and Dybvig (1983) framework is a static, non-monetary model. Accordingly, the type of crisis-prevention intervention they studied–deposit insurance–was non-monetary in nature. In reality, most debt obligations constitute promises to deliver money. And in a run, banks have trouble meeting their short-run nominal obligations. This leads to the question of whether an emergency money-lending facility–a *lender-of-last-resort*–might be designed in a manner to eliminate bank runs.

The lender-of-last-resort function associated with central banking has a long history. Rolnick, Smith and Weber (2000), for example, describe how the Suffolk Bank (a private

[21] At the Friedman rule, monetary equilibrium is not unique because agents are indifferent of how much money they demand since holding money is costless. Accordingly, the real quantity of money is indeterminate but consumption and production are the same in all monetary equilibria. See Lagos and Wright (2005).

bank) acted as a clearinghouse and lender-of-last-resort for the Suffolk Banking System in 19th century New England. Evidently, the Suffolk bank extended loans of specie to its member banks during the crisis of 1837, an action the authors credit with rendering the ensuing recession in the New England area much less severe than in other parts of the country.

Of course, the Suffolk Bank had to rely on reserves of cash in the form of specie. The inelastic supply of specie in commodity money systems is likely what motivated Bagehot (1873) to suggest that central banks lend freely but at "high rates" against good collateral to help stem a bank run. In particular, Bagehot (1873, chapter 7) wrote his first principle as:

> First. That these loans should only be made at a very high rate of interest. This will operate as a heavy fine on unreasonable timidity, and will prevent the greatest number of applications by persons who did not require it. The rate should be raised early in the panic, so that the fine may be paid early; that no one may borrow out of idle precaution without paying well for it; *that the Banking reserve may be protected as far as possible* (italics our own).

That is, the so-called penalty rate of interest was designed, in part, to help protect central bank reserves against depletion. Note that such a constraint is entirely absent if cash takes the form of fiat money instead of specie (or any other commodity). One is led to speculate whether lending freely against good collateral at a *low* interest rate might instead be optimal in a fiat money based system. Indeed, Antinolfi, Huybens and Keister (2001) show that in their model, zero-interest emergency loans are exactly the correct policy for a fiat-based central bank concerned with stemming a liquidity crisis.[22]

We think that a monetary economy is essential for evaluating the efficacy of a central bank lender-of-last-resort policy over other types of interventions designed to stabilize financial markets. The reason for this is based on the simple fact that modern day central banks can costlessly manipulate the supply of base money–which is closely related to the object of redemption in bank deposit contracts. To be sure, the ability to costlessly create money "out of thin air" is often portrayed as a defect, especially for those who fear that governments are too easily seduced by the prospect of inflation finance. But precisely because creating money in this manner is costless, the *threat* of injecting money into the banking system if it is needed can be made *perfectly credible.* This is in contrast to fiscal interventions, which must invariably resort to tax finance in one way or another. While it is possible that a government treasury may raise the money needed for emergency lending, why not have the central bank simply create the needed money at a stroke of a pen?

What is interesting here is that the ability of a central bank lending facility to eliminate bank runs is enhanced greatly if bank deposit contracts are stipulated in *nominal* terms– as they are in reality. A private bank's ability to make good on a nominal promise can be made perfectly credible if it has access to a central bank lending facility. So, let us

[22] Champ, Smith and Williamson (1996) suggest that panics can also be avoided by permitting banks themselves to issue currency.

reconsider our model above and assume that banks and depositors contract in nominal terms, understanding that banks have access to a central bank lending facility that will help banks honor early redemption requests in *any* state of the world. Because the central bank's promise is stated in nominal terms, it is clearly credible. And because banks now do not need to liquidate capital unnecessarily, patient investors know that their deposits will be safe and available for evening withdrawal; hence, a self-fulfilling bank run cannot exist in equilibrium. In this way, depositors may reap the benefits of banking without fearing the losses that are incurred in a banking run.[23] In our model, a central bank lending facility would have the most merit in economies where securities markets are relatively less developed (low η) and the inflation rate is high (say, because money finance constitutes an important source of revenue for the government).

Diamond and Dybvig (1983, § V) mention the use of money creation as a way to secure the resources necessary to finance deposit insurance. But our finding here is more in line with Allen and Gale (1998, § II.D) who show that a central bank intervention consisting of a short-term money loan to the bank can potentially eliminate the deadweight costs of a bank run, though not the the run itself since in their model, the shock driving the run is fundamental and not psychological, as it is here. In our model, demand deposits representing claims against fiat currency can be honored without cost for any price-level, a fact that renders the prospect of self-fulfilling bank runs non-existent under an optimal policy. The notion of nominal debt being conducive to financial stability has been recently explored by Allen, Carletti and Gale (2014). Their model, like ours, relies on some notion of contractual incompleteness. A similar idea underpins the notion of nominal GDP targeting as optimal policy; see, for example, Koenig (2013).

6 Conclusion

Much of what we have learned from the canonical Diamond and Dybvig (1983) and Diamond (1997) models of banking and financial markets appears to remain intact when we extend that framework to a dynamic setting where fiat money plays an essential role. In particular, banks and securities markets remain competing mechanisms for liquidity insurance and bank deposit contracts remain constrained by the availability of financial market alternatives. Moreover, the fragility of banking structures that rely extensively on simple demand deposit liabilities as a form of financing remains a possibility in settings where a central bank is unwilling or unable to operate effectively as a lender of last resort.

Embedding the canonical banking framework in a dynamic monetary model yields the following additional insights. First, easily accessible securities markets together with low-inflation monetary policy can render banking arrangements–together with their associated fragility and potential need for state insurance–superfluous. This result is in contrast to what emerges in a canonical static model, where competitive equilibria are generically sub-optimal. Second, for high-inflation regimes, banking arrangements deliver superior liquidity

[23]Of course, offsetting this benefit is the prospect of moral hazard induced by the lending facility. While our model abstracts from moral hazard, it seems clear that if it was operative, then the obvious trade-off would present itself.

risk-sharing outcomes over securities markets, with the benefit larger in environments where securities markets are missing or are subject to limited participation.

When limited-participation securities markets coexist with banking arrangements, an interesting trade-off emerges between banks and markets as alternative liquidity insurance mechanisms. Conditional on a bank run, a more liquid (high-participation) securities market may result in less *ex post* disruption in individual investment portfolios. However, free access to securities markets may result in less efficient *ex ante* liquidity insurance and, moreover, this is more likely to be the case in high-inflation regimes. Finally, consistent with what has been pointed out by others, the model provides a rationale for nominal debt combined with a central bank lender-of-last-resort facility to promote efficient liquidity insurance and a run-free banking system.

7 References

1. Allen, Franklin and Douglas Gale (2004). Financial Intermediaries and Markets. *Econometrica* 72(4), 1023–1061.

2. Allen, Franklin and Douglas Gale (2007). Understanding Financial Crises. *Oxford University Press*, New York.

3. Allen, Franklin, Carletti, Elena and Douglas Gale (2014). "Money, Financial Stability and Efficiency," *Journal of Economic Theory*, 149(C): 100–127.

4. Andolfatto, David (2010). "Essential Interest-Bearing Money," *Journal of Economic Theory*, 145(4): 1495–1507.

5. Andolfatto, David (2013). "Incentive-Feasible Deflation," *Journal of Monetary Economics*, 60(4): 383–390.

6. Andolfatto, David (2018). "Assessing the Impact of Central Bank Digital Currency on Private Banks," Federal Reserve Bank of St. Louis Working Paper 2018–026B.

7. Antinolfi, Gaetano, Huybens, Elisabeth and Todd Keister (2001). "Monetary Stability and Liquidity Crises: The Role of the Lender of Last Resort," *Journal of Economic Theory*, 99(1-2): 187–219.

8. Bagehot, Walter (1873). *Lombard Street: A Description of the Money Market*, London: Henry S. King and Co.

9. Baumol, William J. (1982). "Contestable Markets: An Uprising in the Theory of Industry Structure," *American Economic Review*, 72(1): 1-15.

10. Bencivenga, Valerie R. and Bruce D. Smith (1991). "Financial Intermediation and Endogenous Growth," *Review of Economic Studies*, 58(2): 195–209.

11. Berentsen, Aleksander, Camera, Gabriele and Christopher Waller (2007). "Money, Credit and Banking," *Journal of Economic Theory* 135(1): 171–195.

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

13. Cooper, Russell and Thomas W. Ross (1998). "Bank runs: liquidity costs and investment distortions", *Journal of Monetary Economics* 41(1): 27–38.

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

15. Diamond, Douglas W. (1997). "Liquidity, Banks, and Markets," *Journal of Political Economy*, 105(5): 928–956.

16. Farhi, Emmanuel, Golosov, Mikhail and Aleh Tsyvinski (2009). "A Theory of Liquidity and Regulation of Financial Intermediation," *Review of Economic Studies*, 76(3): 973–992.

17. Gale, Douglas (1978). "The Core of a Monetary Economy without Trust," *Journal of Economic Theory*, 19(2): 456–491.

18. Geromichalos, Athanasios and Lucas Herrenbrueck (2019). "The Liquidity-Augmented Model of Macroeconomic Aggregates," Simon Fraser University Working Paper.

19. Gorton, Gary (1988). "Banking Panics and Business Cycles," Oxford Economic Papers, 40(4): 751–781.

20. Greenwood, Robin, Hanson, Samuel G. and Jeremy C. Stein (2015). "A Comparative-Advantage Approach to Government Debt Maturity," *Journal of Finance*, 70(4): 1683–1722.

21. Grochulski, Borys and Yuzhe Zhang (2017). "Optimal Liquidity Regulation with Shadow Banking", Federal Reserve Bank of Richmond Working Paper No. 15-12R.

22. Haubrich, Joseph G. (1988). "Optimal Financial Structure in Exchange Economies," *International Economic Review*, 29(2): 217–35.

23. Jacklin, Charles J. (1987). "Demand Deposits, Trading Restrictions, and Risk Sharing," in E. Prescott and N. Wallace (eds), Contractual Arrangements for Intertemporal Trade, Minneapolis, MN: University of Minnesota Press, 26–47.

24. Koenig, Evan F. (2013). "Like a Good Neighbor: Monetary Policy, Financial Stability, and the Distribution of Risk," *International Journal of Central Banking*, 9(2): 57–82.

25. Krishnamurthy, Arvind and Annette Vissing-Jorgensen (2015). "The Impact of Treasury Supply on Financial Sector Lending and Stability," *Journal of Financial Economics*, 118(3): 571–600.

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

27. Loewy, Michael B. (1991). "The Macroeconomic Effects of Bank Runs: An Equilibrium Analysis," *Journal of Financial Intermediation* 1(3): 242–256.

28. Reinhart, Carmen M. and Kenneth S. Rogoff (2009). "The Aftermath of Financial Crises," *American Economic Review*, 99(2): 466–472.

29. Reinhart, Carmen M. and Kenneth S. Rogoff (2014). "Recovery From Financial Crises: Evidence From 100 Episodes," *American Economic Review*, 104(5): 50–55.

30. Rolnick, Arthur J., Smith, Bruce D., and Warren E. Weber (2000). "The Suffolk Bank and the Panic of 1837," Federal Reserve Bank of Minneapolis Quarterly Review 24(2), Spring 2000, 3–13.

31. Skeie, David R. (2008). "Banking with nominal deposits and inside money," *Journal of Financial Intermediation*, 17(4): 562-584.

32. Smith, Bruce D. (2002). "Monetary Policy, Banking Crises, and the Friedman Rule," *American Economic Review*, 92(2): 128-134.

33. Tobin, James (1965). "Money and Economic Growth," *Econometrica*, 33(4): 671–684.

34. von Thadden, Ernst-Ludwig (1997). "The Term-Structure of Investment and the Banks' Insurance Function," *European Economic Review*, 41(7): 1355–1374.

A Proofs

Proof of Proposition 1. Consider the problem of an investor in a securities market with limited participation: (23) subject to (12)–(22). Although omitted in the main body, here, we explicitly consider the non-negativity constraints on x and k. Note that k_1 and k_2 can, in principle, take negative values; $k_1 < 0$ ($k_2 < 0$) means the impatient (patient) investor is buying (selling) securities. Hence, we need to add $k + k_2 \geq 0$ as a constraint for patient investors with market access. The following results simplify the analysis: (17) and (19) imply (18), so that constraint (18) can be ignored; and $V'(z) = 1$.

Let $\pi\rho\psi_1$ and $(1-\pi)\eta\rho\psi_2$ be the Lagrange multipliers associated with the no-securities-shorting constraint (14) and $k+k_2 \geq 0$, respectively. Let $\pi\lambda_1$, $(1-\pi)\eta\lambda_2$ and $(1-\pi)(1-\eta)\hat{\lambda}_2$ denote the Lagrange multipliers associated with constraints (15), (19) and (22), respectively. Let $\pi\beta\zeta_1$, $(1-\pi)\eta\beta\zeta_2$ and $(1-\pi)(1-\eta)\beta\hat{\zeta}_2$ denote the Lagrange multipliers associated with the non-negativity constraints (13), (17) and (21), respectively. Let ξ_x and ξ_k be the Lagrange multipliers associated with the non-negativity constraints $x \geq 0$, and $k \geq 0$, respectively. The first-order conditions imply:

$$
\begin{aligned}
\pi[u'(c_1) + \lambda_1] + (1-\pi)\eta[u'(c_2) + \lambda_2] + (1-\pi)(1-\eta)[u'(\hat{c}_2) + \hat{\lambda}_2] &= (\mu/\beta)(1 - \xi_x) && (63)\\
\pi[ru'(c_1) + \rho\psi_1] + (1-\pi)\eta[Ru'(c_2) + \rho\psi_2] + (1-\pi)(1-\eta)Ru'(\hat{c}_2) &= 1 - \xi_k && (64)\\
(\rho - r)u'(c_1) &= \rho[\psi_1 - \lambda_1] && (65)\\
\eta(R-\rho)u'(c_2) &= \eta\rho[\lambda_2 - \psi_2] && (66)\\
u'(c_1) - 1 + \lambda_1 &= \zeta_1 && (67)\\
\eta[u'(c_2) - 1 + \lambda_2] &= \eta\zeta_2 && (68)\\
(1-\eta)[u'(\hat{c}_2) - 1 + \hat{\lambda}_2] &= (1-\eta)\hat{\zeta}_2 && (69)
\end{aligned}
$$

Market clearing implies:

$$\pi k_1 = (1-\pi)\eta k_2 \tag{70}$$

Let $0 < \eta \leq 1$. Zero consumption contradicts optimality by $u'(0) = \infty$. Hence, we cannot have $x = k = 0$. From (65) and (66), $\rho = 0$ implies $-ru'(c_1) = Ru'(c_2) = 0$, a contradiction with u strictly increasing; hence, $\rho > 0$. Combining (63)–(66) we obtain:

$$(1-\pi)(1-\eta)[(R/\rho - 1)u'(\hat{c}_2) - \hat{\lambda}_2] = (1/\rho)(1-\xi_k) - (\mu/\beta)(1-\xi_x) \tag{71}$$

From (63)–(64) we have $0 \leq \xi_x < 1$ and $0 \leq \xi_k < 1$.

Let $\eta = 1$. We guess and verify $x > 0$ and $k > 0$. Since $\xi_x = \xi_k = 0$, (71) implies $\rho = \beta/\mu$. By (65) and (66) we get $\psi_1 > 0$ and $\lambda_2 > 0$. Thus, $k = k_1$ and $(\beta/\mu)x - \rho k_2 + z_2^+ = 0$. The latter condition implies $c_2 = R(k + k_2)$. We thus verify $k > 0$, since otherwise $k = k_1 = k_2 = c_2 = 0$. Given $k_1 = k > 0$, we also verify $x > 0$, since otherwise $k_1 = 0$. By (70) $k_1 > 0$ implies $k_2 > 0$ and so $\psi_2 = 0$. Given $k_1 = k$, $c_1 = (\beta/\mu)x + \rho k_1 + z_1^+ > 0$ and so, $\lambda_1 = 0$.

Now let $0 < \eta < 1$. There are five regions to consider: $0 < \rho < r$; $\rho = r$; $r < \rho < R$; $\rho = R$; and $\rho > R$. We show that only $0 < \rho < \beta/\mu$, $x > 0$ and $k > 0$ can be an equilibrium.

1) Suppose that $0 < \rho < r$, which by (65) and (66) implies $\lambda_1 > 0$ and $\lambda_2 > 0$. Then, $(\beta/\mu)x + \rho k_1 - \beta z_1^+ = (\beta/\mu)x - \rho k_2 - \beta z_2^+ = 0$. Hence, $c_1 = r(k - k_1) > 0$ and

$c_2 = R(k+k_2) > 0$, which implies $k > 0$, $k-k_1 > 0$ and $k+k_2 > 0$. Thus, $\xi_k = \psi_1 = \psi_2 = 0$. If $x = 0$ then $k_1 = k_2 = 0$ and so, $c_1 = rk < c_2 = \hat{c}_2 = Rk$. Condition (64) implies

$$\pi r u'(rk) + (1-\pi) R u'(Rk) = 1$$

By [A1] and $r < 1 < R$, $u'(rk) > 1$ and $u'(Rk) < 1$. By (67), (68), $\lambda_1 > 0$ and $\lambda_2 > 0$, we get $\zeta_1 > 0$ and $z_1^+ = z_2^+ = \zeta_2 = 0$. By $\zeta_2 = 0$, (66) and (68), $(R/\rho)u'(Rk) = 1$. Since $\hat{c}_2 = Rk$, (69) implies $\hat{\lambda}_2 > 0$ and so, $\hat{z}_2^+ = \hat{\zeta}_2 = 0$. Then, $\hat{\lambda}_2 = 1 - \rho/R$ and by (71)

$$(1-\pi)(1-\eta)[(R/\rho)u'(Rk) - 1] = 1/\rho - (\mu/\beta)(1-\xi_x)$$

Since $(R/\rho)u'(Rk) = 1$ we obtain $1/\rho = (\mu/\beta)(1-\xi_x)$. By [A2] $r(\mu/\beta) < 1$ and recall that $0 \leq \xi_x < 1$. Hence, $1 = \rho(\mu/\beta)(1-\xi_x) < r(\mu/\beta)(1-\xi_x) < 1$, a contradiction. Hence, $x > 0$.

2) Suppose that $\rho = r$. By (65) $\psi_1 = \lambda_1 \geq 0$. If $\psi_1 > 0$ then $k = k_1$ and since $\lambda_1 > 0$, $(\beta/\mu)x + \rho k_1 - \beta z_1^+ = 0$ and so, $c_1 = r(k-k_1) = 0$, a contradiction. Hence, $\psi_1 = \lambda_1 = 0$. By (66) $\lambda_2 > 0$ and so, $(\beta/\mu)x - \rho k_2 - \beta z_2^+ = 0$ which implies $c_2 = R(k+k_2)$. Thus, $k > 0$ and $\xi_k = \psi_2 = 0$, since otherwise $c_2 = 0$. If $x = 0$ then $k_1 = k_2 = 0$ and $c_1 = rk < c_2 = \hat{c}_2 = Rk$. This case leads to the same contradiction as in 1), $1 < r(\mu/\beta)(1-\xi_x) < 1$. Thus, $x > 0$.

3) Suppose $r < \rho < R$. By (65) and (66) we get $\psi_1 > 0$ and $\lambda_2 > 0$, and so, $k = k_1$ and $(\beta/\mu)x - \rho k_2 - \beta z_2^+ = 0$. Thus, $c_1 = (\beta/\mu)x + \rho k_1 - \beta z_1^+ > 0$, which implies $\lambda_1 = 0$, and $c_2 = R(k+k_2) > 0$, which implies $k > 0$ and $\xi_k = \psi_2 = 0$. Since $k = k_1 > 0$, $x > 0$ and $\xi_x = 0$. By (71) we get

$$(1-\pi)(1-\eta)[(R/\rho - 1)u'(\hat{c}_2) - \hat{\lambda}_2] = 1/\rho - \mu/\beta$$

If $\hat{\lambda}_2 = 0$ then $\rho < \beta/\mu$. If $\hat{\lambda}_2 > 0$ then $\hat{c}_2 = Rk < c_2 = R(k+k_2)$ and $(\beta/\mu)x = \beta\hat{z}_2^+ > 0$ and so, $\hat{\zeta}_2 = 0$. Since $\lambda_2 > 0$, $\psi_2 = 0$ and $\zeta_2 \geq 0$, (66) and (68) imply $(R/\rho)u'(c_2) - 1 \geq 0$. Since $c_2 > \hat{c}_2$, $(R/\rho)u'(\hat{c}_2) - 1 > 0$. Since $\hat{\zeta}_2 = 0$, (69) implies $\hat{\lambda}_2 = 1 - u'(\hat{c}_2)$ and so, $(R/\rho - 1)u'(\hat{c}_2) - \hat{\lambda}_2 = (R/\rho)u'(\hat{c}_2) - 1 > 0$. Thus, $\rho < \beta/\mu$.

4) Suppose that $\rho = R$. By (66), $\lambda_2 = \psi_2 \geq 0$. If $\lambda_2 = \psi_2 > 0$ then $(\beta/\mu)x - \rho k_2 - \beta z_2^+ = 0$ and thus, $c_2 = R(k+k_2) = 0$, a contradiction. Hence, $\lambda_2 = \psi_2 = 0$. By (65), $\psi_1 > 0$ and so, $k = k_1$. Hence, $c_1 = (\beta/\mu)x + \rho k_1 - \beta z_1^+ > 0$, which implies $\lambda_1 = 0$. If $x = 0$ then $k = k_1 = 0$, a contradiction with positive consumption. Hence, $x > 0$ and $\xi_x = 0$. By (71), $0 \leq \xi_k < 1$ and [A2] we get $(1-\pi)(1-\eta)\hat{\lambda}_2 = \mu/\beta - (1-\xi_k)/R > 0$. Then, $(\beta/\mu)x = \beta\hat{z}_2^+ > 0$ and so, $\hat{c}_2 = Rk$ and $\hat{\zeta}_2 = 0$. Hence, $k > 0$, and $\xi_k = 0$. Since $k_1 = k > 0$, (70) implies $k_2 > 0$ and thus, $c_2 > \hat{c}_2$. By (69), $\hat{\lambda}_2 = 1 - u'(\hat{c}_2) > 0$ and so, $u'(\hat{c}_2) < 1$. Since $c_2 > \hat{c}_2$, $u'(c_2) < u'(\hat{c}_2) < 1$, which given $\lambda_2 = 0$ implies $\zeta_2 < 0$ by (68), a contradiction.

5) Suppose that $\rho > R$, which by (65) and (66) implies $\psi_1 > 0$ and $\psi_2 > 0$. Then, $k_1 = -k_2 = k$ and by (70) $k = k_1 = k_2 = 0$. Then, $\lambda_1 = \lambda_2 = \hat{\lambda}_2 = \xi_x = 0$ and $x > 0$ (otherwise consumption in all states is zero). We guess and verify that $z_1^+ = z_2^+ = \hat{z}_2^+ = 0$, which implies $c_1 = c_2 = \hat{c}_2 = c = (\beta/\mu)x > 0$. Using these results (63) simplifies to $u'(c) = \mu/\beta > 1$. By (67)–(69) we verify $\zeta_1 > 0$, $\zeta_2 > 0$ and $\hat{\zeta}_2 > 0$, and hence, $z_1^+ = z_2^+ = \hat{z}_2^+ = 0$. Using (63)–(66) we get: $[\pi + (1-\pi)\eta]\rho + (1-\pi)(1-\eta)R = (\beta/\mu)(1-\xi_k)$. By [A2], $\rho > R > 1$ and $\xi_k \geq 0$, this condition implies $1 < R < (\beta/\mu)(1-\xi_k) < 1$, a contradiction. ■

Proof of Lemma 1. Let $\eta = 1$. By Proposition 1, $\rho = \beta/\mu$, $\psi_1 > 1$, $\lambda_1 = 0$ and $\lambda_2 > 0$. Hence, $k_1 = k$ and using (70) $k_2 = \pi k/(1-\pi)$ and $\beta z_2^+ = (\beta/\mu)[x - \pi k/(1-\pi)]$. The remaining first-order conditions are:

$$\begin{aligned} \pi u'(c_1) + (1-\pi) R(\mu/\beta) u'(c_2) &= \mu/\beta && (72)\\ u'(c_1) - 1 &= \zeta_1 && (73)\\ R(\mu/\beta) u'(c_2) - 1 &= \zeta_2 && (74) \end{aligned}$$

Since $\beta z_1^+ = (\beta/\mu)(x+k) - c_1$ and $\beta z_2^+ = (\beta/\mu)[x + \pi k/(1-\pi)]$ the complementary slackness conditions are

$$\begin{aligned} \zeta_1 [(\beta/\mu)(x+k) - c_1] &= 0 && (75)\\ \zeta_2 (\beta/\mu)\left(x - \frac{\pi k}{1-\pi}\right) &= 0 && (76) \end{aligned}$$

There are four possible cases for the values of the multipliers ζ_1 and ζ_2.

1) Assume $\zeta_1 = \zeta_2 = 0$: both impatient and patient investors save money across period. From (72)–(74), we get $1 = \mu/\beta$, which contradicts [A2]. This case is only possible under the Friedman rule.

2) Assume $\zeta_1 > 0$ and $\zeta_2 = 0$: only patient investors save money across periods. From (72)–(74), we obtain

$$u'(c_1) = \frac{\mu/\beta - (1-\pi)}{\pi} \text{ and } R(\mu/\beta) u'(c_2) = 1, \tag{77}$$

and from (32), (75), and (76) we obtain

$$c_1 = (\beta/\mu)(x+k), \; c_2 = Rk/(1-\pi) \text{ and } (1-\pi) x \geq \pi k. \tag{78}$$

Existence of case 2 requires that $u'(c_1) \geq 1$ and $(1-\pi)x \geq \pi k$. It is easy to show that $u'(c_1) = \frac{\mu/\beta-(1-\pi)}{\pi} > 1$, unless $\mu = \beta$. Rewrite $(1-\pi)x \geq \pi k$ to get $(1-\pi)(x+k) \geq k$. Use (78) to get

$$c_1 R(\mu/\beta) \geq c_2$$

Use (77) to get

$$u'(c_1)c_1 \geq \frac{\mu/\beta - (1-\pi)}{\pi} u'(c_2)c_2$$

This case requires $u'(c)c$ to be decreasing as we assume with [A1].

3) Assume $\zeta_1 = 0$ and $\zeta_2 > 0$: only impatient investors save money across periods. From (72)–(74), we obtain

$$u'(c_1) = 1 \text{ and } R(\mu/\beta) u'(c_2) = \frac{\mu/\beta - \pi}{(1-\pi)} \tag{79}$$

and from (32), (75) and (76) we obtain $c_1 \leq (\beta/\mu)(x+k)$, $c_2 = Rk/(1-\pi)$ and $(1-\pi)x = \pi k$. Existence requires $R(\mu/\beta) u'(c_2) \geq 1$ and $c_1 \leq (\beta/\mu)x/\pi$. It is easy to show that

$R\,(\mu/\beta)\,u'(c_2) = R\,(\mu/\beta)\,\frac{1-(\beta/\mu)\pi}{(1-\pi)R} > 1$, unless $\mu = \beta$. Rewrite $c_1 \leq (\beta/\mu)\,(x+k)$ to get $c_1 \leq (\beta/\mu)k/(1-\pi)$. Rewrite again to get $c_1 R\,(\mu/\beta) \leq c_2$, which using (79) implies

$$u'(c_1)c_1 < u'(c_1)c_1 \left[\frac{\mu/\beta - \pi}{1-\pi}\right] \leq u'(c_2)c_2$$

This case requires $u'(c)c$ to be increasing which [A1] precludes.

4) Assume $\zeta_1 > 0$ and $\zeta_2 > 0$: no money savings across periods. From (72)–(74), we obtain

$$\pi u'(c_1) + (1-\pi)\,R\,(\mu/\beta)\,u'(c_2) = \mu/\beta \tag{80}$$

and from (32), (75) and (76) we obtain

$$c_1 = (\beta/\mu)\,(x+k)\,,\;\; c_2 = Rk/(1-\pi) \text{ and } (1-\pi)\,x = \pi k \tag{81}$$

Note that $c_2/c_1 = R\,(\mu/\beta)$. This case requires $u'(c_1) \geq 1$ and $R\,(\mu/\beta)\,u'(c_2) \geq 1$. ■

Proof of Proposition 2. We begin by showing that $Rc_1^* > c_2^*$. By [A1], $R > 1$ and (6) we have $Rc_1^* u'(Rc_1^*) < c_1^* u'(c_1^*) = c_1^* R u'(c_2^*)$, which implies $u'(Rc_1^*) < u'(c_2^*)$ and so, $Rc_1^* > c_2^*$.

Evaluating (42) at $\mu = \beta$ and (c_1^*, c_2^*), we get $z_2^+ = c_1^* - c_2^*/R > 0$, i.e., patient investors carry strictly positive amounts of money across periods if the first-best is implemented at the Friedman rule. Finally, from (6) we have $u'(c_1^*) = Ru'(c_2^*) = 1$ which satisfy (40) and (41) when $\mu = \beta$. Hence, the Friedman rule implements the first-best. ■

Proof of Proposition 3. Follows from (44) and (45). ■

Proof of Lemma 2. Implementing the first-best allocation in a securities market equilibrium with limited participation $(0 < \eta < 1)$ implies $c_1 = c_1^*$ and $c_2 = \hat{c}_2 = c_2^*$. Recall that $u'(c_1^*) = Ru'(c_2^*) = 1$, with $x^* > 0$, $k^* > 0$. At the first-best allocation, $\xi_x = \xi_k = \psi_2 = 0$. Combining (64) and (65) implies $\rho = 1$ and so, $\psi_1 > 0$, $\lambda_1 = 0$, $\lambda_2 > 0$. By (69), $1/R - 1 + \hat{\lambda}_2 = \hat{\zeta}_2 > 0$ and so, $\hat{\lambda}_2 = 0$. Condition (63) simplifies to

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

a contradiction by [A2], $R > 1$ and $0 < \eta < 1$: as $\mu \searrow \beta$, the right-hand side converges to 1, while the left-hand side remains below 1. Thus, the first-best allocation cannot be implemented even at the Friedman rule. ■

Proof of Proposition 4. Follows from (54) and (55). ■

Proof of Proposition 5. First, note that $V(z) = V(0) + z$ and $W(z) = W(0) + z$, so that $V(z) - W(z) = V(0) - W(0)$ for all $z \geq 0$. Second, note that a bank maximizes the *ex ante* welfare of an investor and can always implement the securities equilibrium allocation, as feasibility is the only relevant constraint in the banking problem. Furthermore, given (77) (for the case when patient investors save cash) and (81) (for the case when patient investors do not save cash), $c_1 < c_2$ in a securities equilibrium; hence, (c_1, c_2) is an incentive compatible allocation for the bank.

We will now show that when $\mu > \beta$ the bank, despite being able to implement the securities equilibrium with $\eta = 1$ as shown above, chooses a different allocation. Thus, it must be that the bank allocation provides higher *ex ante* utility for the investor, i.e., $W(z) > V(z)$. Suppose the securities equilibrium and the bank choose the same allocation. Then, both choose $c_2 = c_2^*$ and so $k = k^*$. By (66) and (68), $\rho = \beta/\mu < 1$ and $c_2 = c_2^*$ imply $1/\rho - 1 = \zeta_2 > 0$ and so $z_2^+ = 0$. Then (39) applies in a securities market equilibrium and so, $(\mu/\beta)c_1 = c_2^*/R$. Given $c_1 < c_2$, assumption [A1] implies $u'(c_1)c_1 > u'(c_2)c_2$. Multiply both sides by $R(\beta/\mu)$ to obtain $R(\beta/\mu)u'(c_1)c_1 > R(\beta/\mu)u'(c_2)c_2$. Using $(\beta/\mu)u'(c_1) = Ru'(c_2) = 1$, this simplifies to $Rc_1 > (\beta/\mu)c_2$ and so $(\mu/\beta)c_1 > c_2/R$, which contradicts $(\mu/\beta)c_1 = c_2^*/R$. Thus, banks and markets implement different allocations. ■

Proof of Proposition 6. Assuming [A2], conditions (52)–(53) imply $c_1 < c_1^* < c_2 = c_2^*$. A patient depositor withdrawing his deposit in the afternoon and using the cash to buy capital in the securities market at price ρ gets evening consumption equal to $\tilde{c}_2 = (R/\rho)c_1$. Let $\bar{\rho} \equiv R(c_1/c_2^*)$. If $\rho < \bar{\rho}$ then $\tilde{c}_2 > (R/\bar{\rho})c_1 = c_2^*$. Hence, the decision to withdraw in the afternoon is optimal. If $\rho \geq \bar{\rho}$ then $\tilde{c}_2 \leq (R/\bar{\rho})c_1 = c_2^*$ and so, there are no incentives to withdraw in the afternoon.

Now we show $\beta/\mu < \bar{\rho} < R$. The latter inequality follows from $c_1 < c_2^*$. The former inequality requires [A1]. This assumption implies $u'(c_1)c_1 > u'(c_2^*)c_2^*$. Multiply both sides by $R(\beta/\mu)$ to obtain $R(\beta/\mu)u'(c_1)c_1 > R(\beta/\mu)u'(c_2^*)c_2^*$. Again using (52)–(53) we get $(\beta/\mu)u'(c_1) = Ru'(c_2^*) = 1$, which simplifies the previous expression to $Rc_1 > (\beta/\mu)c_2^*$ and so $R(c_1/c_2^*) > \beta/\mu$. ■

Proof Proposition 7. Condition (61) implies $c_2 = c_2^* > c_1^*$. Let $\tilde{\pi} \equiv \pi/[\pi + (1-\pi)\tilde{\eta}]$ and rewrite (60) as

$$\tilde{\pi}u'(c_1) + (1-\tilde{\pi})(R/\rho)u'(\tilde{c}_2) = \mu/\beta \tag{82}$$

If $\tilde{\zeta} = 0$ then (62) implies $(R/\rho)u'(\tilde{c}_2) = 1$ and so, by [A2], $u'(c_1) > \mu/\beta$. If $\tilde{\zeta} > 0$ then (57) implies $(R/\rho)c_1 = \tilde{c}_2$. Thus, (82) implies $\tilde{\pi}u'(c_1) + (1-\tilde{\pi})(R/\rho)u'((R/\rho)c_1) = \mu/\beta$. By [A1] and $\rho < R$, $(R/\rho)u'((R/\rho)c_1) < u'(c_1)$. Hence, $\tilde{\pi}u'(c_1) + (1-\tilde{\pi})(R/\rho)u'((R/\rho)c_1) < u'(c_1)$ and so, $u'(c_1) > \mu/\beta$ in this case as well.

We have $u'(c_1) > \mu/\beta > 1 = u'(c_1^*)$ and from (82), $(R/\rho)u'(\tilde{c}_2) < \mu/\beta$. The first set of inequalities imply $c_1 < c_1^* < c_2^* = c_2$ and so, the incentive-compatibility constraint $u(c_2) \geq u(c_1)$ is satisfied. Furthermore, let c_1^B solve (52); then, $c_1 < c_1^B < c_1^*$. The second inequality implies $Ru'(\tilde{c}_2) < \rho(\mu/\beta)$. By [A3], $\rho(\mu/\beta) \leq 1$ and so, $Ru'(\tilde{c}_2) < 1 = Ru'(c_2^*)$. Hence, $\tilde{c}_2 > c_2^*$. Given $c_1 - (\rho/R)\tilde{c}_2 \geq 0$ from (57), $\tilde{c}_2 > c_2^*$ implies the incentive-compatibility constraint $u(\tilde{c}_2) + c_1 - (\rho/R)\tilde{c}_2 \geq u(c_2)$ is satisfied. ■

Proof of Proposition 8. By Proposition 7 the incentive constraints are slack. Thus, totally differentiating $W(z)$ with respect to μ yields:

$$dW/d\mu = -(1/\mu)\left\{\pi u'(c_1)c_1 + (1-\pi)\tilde{\eta}(1+\tilde{\zeta})[c_1 + (d\rho/d\mu)(\mu/R)\tilde{c}_2]\right\}$$

where we used the fact that $(\beta/\mu)\,x = [\pi + (1-\pi)\,\tilde{\eta}]\,c_1$ and that ρ is a function of μ.

If $c_1 \geq -(d\rho/d\mu)(\mu/R)\tilde{c}_2$ then $dW/d\mu < 0$. From constraint (57) we have $c_1 \geq (\rho/R)\tilde{c}_2$. Thus, it is sufficient to require $-(d\rho/d\mu)\mu \leq \rho$, equivalently, $-(d\rho/d\mu)(\mu/\rho) \leq 1$, to obtain $dW/d\mu < 0$. ■

Proof of Proposition 9. Consider the unbanked investors first. By Proposition 2 and Lemma 2, the first best allocation is implemented in a securities market equilibrium if and only if $\eta = 1$ and $\mu = \beta$. From Proposition 2 we know that first-best implementation requires positive money savings by patient investors.

Consider now the banking problem characterized by (57)–(62). Suppose that $\mu > \beta$. By Proposition 7, $c_1 < c_1^*$ and $\tilde{c}_2 > c_2^*$; thus, the first-best cannot be implemented.

Suppose that $\mu = \beta$. Then condition (60) can be rearranged as

$$\pi[u'(c_1) - 1] = (1-\pi)\tilde{\eta}[1 - (R/\rho)u'(\tilde{c}_2)]$$

At the first-best, $c_1 = c_1^*$ and hence, both sides of the expression above must be equal to zero. This implies $(R/\rho)u'(\tilde{c}_2) = 1$. We get $\tilde{c}_2 = c_2^*$ only if $\rho = 1$. From Proposition 1, $\rho < \beta/\mu$ when $\eta \in (0,1)$. Thus, the first-best cannot be achieved for any $\eta \in (0,1)$.

The remaining case is $\mu = \beta$ and $\eta = 1$. By (26), $\rho = 1$. Conditions (60)–(62) become

$$\pi u'(c_1) + (1-\pi)\tilde{\eta}Ru'(\tilde{c}_2) = \pi + (1-\pi)\tilde{\eta} \quad (83)$$
$$Ru'(c_2) = 1 \quad (84)$$
$$Ru'(\tilde{c}_2) = 1 + \tilde{\zeta} \quad (85)$$

If $\tilde{\zeta} > 0$, then (84) and (85) imply $\tilde{c}_2 < c_2$, a contradiction with Proposition 7.

If $\tilde{\zeta} = 0$ then (83)–(85) imply $u'(c_1) = Ru'(c_2) = Ru'(\tilde{c}_2) = 1$, so that $c_1 = c_1^*$ and $\tilde{c}_2 = c_2 = c_2^*$. All incentive constraints are satisfied since $\tilde{c}_2 = c_2 > c_1$. From (57), strictly positive money savings requires that $c_1 > \tilde{c}_2/R$. That is, we need $u'(c_1)c_1 > u'(\tilde{c}_2)\tilde{c}_2$, which holds given [A1] and $c_1 < \tilde{c}_2$. Hence, when $\mu = \beta$ and $\eta = 1$ the first-best allocation is feasible, incentive-compatible and requires positive money savings by patient depositors with market access. ■

Proof of Lemma 3. (i) Assume $0 < \tilde{\eta} < 1$. Condition (57) implies $\tilde{c}_2 \leq (R/\rho)c_1$. Since $c_2 = c_2^* < \tilde{c}_2$ by Proposition 7, we get $c_2 < (R/\rho)c_1$. From the bank's budget constraints, we have $(\beta/\mu)x = [\pi + (1-\pi)\tilde{\eta}]c_1$ and $Rk = (1-\pi)(1-\tilde{\eta})c_2$. Thus, $(\beta/\mu)x + \rho k = [\pi + (1-\pi)\tilde{\eta}]c_1 + (1-\pi)(1-\tilde{\eta})(\rho/R)c_2$. Since $c_2 < (R/\rho)c_1$, we get $(\beta/\mu)x + \rho k < [\pi + (1-\pi)\tilde{\eta}]c_1 + (1-\pi)(1-\tilde{\eta})c_1 = c_1$.

(ii) From the bank's budget constraints, when $\tilde{\eta} = 1$, $(\beta/\mu)x = c_1$ and $k = 0$. ■

Proof of Proposition 10.

A bank liquidating its capital in the securities market hands out $(\beta/\mu)x + \rho k$ to each depositor. By Lemma 3, $(\beta/\mu)x + \rho k < c_1$ and by Proposition 7, $c_1 < c_2$. Hence, both impatient depositors and patient depositors without market access obtain the same allocation, $c_1^l = c_2^l = (\beta/\mu)x + \rho k < c_1 < c_2$.

A patient depositor with market access can sell c_1^l units of cash for capital and obtain a maximum consumption of $(R/\rho)c_1^l$. If $(R/\rho)u'(c_1^l) \leq 1$ then $(R/\rho)u'((R/\rho)c_1) < 1$. Thus, by (57) and (62), $(R/\rho)u'(\tilde{c}_2) = 1$ and $\tilde{c}_2 \leq (R/\rho)c_1^l < (R/\rho)c_1$. Since $\tilde{c}_2^l \leq (R/\rho)c_1^l$, $\tilde{c}_2^l = \tilde{c}_2$ is feasible and optimal.

If $(R/\rho)u'(c_1^l) > 1$ then $\tilde{c}_2^l = (R/\rho)c_1^l$. If (57) is satisfied with equality, then $\tilde{c}_2^l = (R/\rho)c_1 > (R/\rho)c_1^l = \tilde{c}_2^l$. If (57) is satisfied with strict inequality, then $(R/\rho)u'(\tilde{c}_2) = 1$ and so, $\tilde{c}_2 > (R/\rho)c_1^l = \tilde{c}_2^l$. ■

Proof of Proposition 11. Each depositor receives $(\beta/\mu)x$ units of cash and k claims on capital. Thus, impatient depositors obtain $c_1^h = (\beta/\mu)x + \rho k = c_1^l$. Patient depositors with market access can get a maximum consumption of $(R/\rho)(\beta/\mu)x + Rk = (R/\rho)c_1^h$, same as before. Thus, $\tilde{c}_2^h = \tilde{c}_2^l$. Patient depositors without market access buy afternoon goods with cash and hold their claims on capital until the evening; hence, $c_2^h = (\beta/\mu)x + Rk > c_2^l$. Using the budget constraints of the bank we get $(\beta/\mu)x + Rk = [\pi + (1-\pi)\tilde{\eta}]c_1 + (1-\pi)(1-\tilde{\eta})c_2$. Since $c_1 < c_2$ by Proposition 7, $c_2^h = (\beta/\mu)x + Rk < c_2$. ■

B Upper bound on securities price

We demonstrate here that workers accept only money (and not securities) in the afternoon. They do so because money can be used to buy output the next morning (something workers value), whereas securities deliver output only in the evening (when workers have no desire to consume). This fact alone, however, does not rule out the possibility that workers might want to buy (or short) securities in the morning and sell (or repurchase) them in the afternoon to exploit an arbitrage opportunity. We can rule out a shorting strategy by appealing to the assumed lack of commitment and the fact that workers have no incentive to purchase securities in the afternoon. There is still a possibility that workers may want to buy securities in the morning. In what follows, we explain what rules out this possibility.

Consider the problem of a worker that wants to buy capital in the morning. Since capital depreciates in the evening and the worker only enjoys morning consumption, the only option is to exchange the capital acquired in the morning for cash in the afternoon securities market. The relative price of worker consumption and investor output in the morning is equal to 1. Thus, every unit of capital acquired by the worker costs 1 units of utility. Capital is exchanged for money in the afternoon at price q_t, which can be used in the following morning to purchase consumption at price p_{t+1}^m. Assume the worker can access the afternoon securities market with probability $\pi_w \in [0, 1]$.

The problem of the worker buying capital in the morning is:

$$\max_{k \geq 0} \; -k + \pi_w \beta (q_t / p_{t+1}^m) k$$

Note that the worker cannot go short on capital, since he can neither produce it nor commit to repay debts. Since $\rho \equiv q_t / p_t^a$ and $p_t^a / p_{t+1}^m = (p_t^a / p_t^m)(p_t^m / p_{t+1}^m) = (\mu/\beta)(1/\mu) = 1/\beta$, we have that $\rho = \beta(q_t / p_{t+1}^m)$. Thus, we can rewrite the problem of the worker as

$$\max_{k \geq 0} \; -k + \pi_w \rho k$$

An interior solution is given by $\rho = 1/\pi_w$, in which case workers are indifferent between selling and purchasing capital. If $\rho < 1/\pi_w$ then $k \geq 0$ binds and workers do not buy any capital. If $\rho > 1/\pi_w$ then workers demand an infinite amount of capital. Hence, in a monetary equilibrium $\rho \leq 1/\pi_w$. If workers have perfect market access, $\pi_w = 1$, then $\rho \leq 1$. If workers have limited market access, $\pi_w < 1$, then the upper bound on ρ is larger than 1; moreover, the upper bound converges to infinity as market access for workers goes to zero.

The upper bound on the securities price is not operative in our analysis, since $\rho \leq \beta/\mu \leq 1$. It could play a role in environments where impatient investors have limited market access, in which case ρ could exceed β/μ and rise as high as R.

C Alternative bank contract

In Section 4, we study a bank contract in which the bank does not try to prevent patient depositors with market access from withdrawing in the afternoon. Below, we formally verify the validity of this approach, showing that the contract we study is more general than one in which, instead, the bank provides incentives to prevent patient depositors with market access from withdrawing in the afternoon.

Consider a bank contract that assigns cash in the amount of c_1 for depositors withdrawing in the afternoon. A patient depositor with market access that withdraws in the afternoon would obtain c_1 in cash, which can (at least in part) be exchanged for capital in the securities market, delivering $\tilde{c}_2$ in the evening, and save the remainder cash (if any) $\beta\tilde{z}_2^+$ for the following period. Given that $\beta\tilde{z}_2^+ = c_1 - (\rho/R)\tilde{c}_2 \geq 0$, the value of this deviation strategy is determined by the following decision problem:

$$D(c_1) \equiv \max_{\tilde{c}_2} \; u(\tilde{c}_2) + c_1 - (\rho/R)\tilde{c}_2 + \tilde{W}(0)$$

subject to $c_1 - (\rho/R)\tilde{c}_2 \geq 0$ and where $\tilde{W}$ is the value of reentering a deposit contract next period, which as usual is linear in cash holdings.

Let $\tilde{\lambda}$ be the Lagrange multiplier associated with the non-negativity constraint on cash holdings. Then, the first-order condition implies

$$\tilde{\lambda} = (R/\rho)u'(\tilde{c}_2) - 1$$

Hence, either $Ru'(\tilde{c}_2) = \rho$ with $\tilde{\lambda} = 0$ and $\beta\tilde{z}_2^+ \geq 0$; or $\tilde{c}_2 = (R/\rho)c_1$ with $\tilde{\lambda} > 0$ and $\beta\tilde{z}_2^+ = 0$.

Let $d(c_1) \equiv D(c_1) - \tilde{W}(0) = u(\tilde{c}_2) + c_1 - (\rho/R)\tilde{c}_2$. Then $d'(c_1) = 1 + \tilde{\lambda} > 0$, i.e., the value of deviating is strictly increasing in c_1.

Consider then a bank contract structured to prevent afternoon redemptions by patient depositors with market access. Such a contract implies the following incentive-compatibility constraint

$$u(c_2) + \beta z_2^+ - d(c_1) \geq 0 \tag{86}$$

where we allow the bank to leave some unspent cash for evening withdrawals, as this may help provide incentives to patient depositors with market access.

The problem of the bank can be written as

$$\tilde{W}(z) \equiv \max_{x,k,z_2^+} \; z - x - k + \pi u(c_1) + (1-\pi)[u(c_2) + \beta z_2^+] + \tilde{W}(0) \tag{ALT}$$

subject to (86) and

$$\begin{aligned} u(c_2) + \beta z_2^+ - u(c_1) &\geq 0 \quad (87) \\ z_2^+ &\geq 0 \quad (88) \end{aligned}$$

where

$$\begin{aligned} c_1 &= [(\beta/\mu)x - (1-\pi)\beta z_2^+]/\pi \\ c_2 &= Rk/(1-\pi) \end{aligned}$$

If (86) were slack then we would recover the bank contract from Section 3, where setting $z_2^+ = 0$ is optimal. However, Proposition 6 showed that this contract is not incentive-compatible for patient depositors with market access. Hence, (86) will bind. In contrast, we anticipate that (87) will not bind. The non-negativity constraint (88) may or may not bind. The lemma below formally characterizes the properties of this alternative contract.

Lemma 4 *In the bank contract given by problem* (ALT)*: (i) the incentive constraint* (87) *is slack; (ii) the incentive constraint* (86) *binds; and (iii)* $c_1 < c_1^B < c_1^* < c_2^B = c_2^* < c_2$*, where* (c_1^B, c_2^B) *solves* (52)–(53).

Proof. (i) We begin by showing that $d(c_1) > u(c_1)$ and so, (86) implies (87). We only need to consider the case when the liquidity constraint in the deviation problem binds, as if it were slack then $d(c_1)$ would be even larger. Hence, suppose that $\tilde{c}_2 = (R/\rho)c_1$, which implies $d(c_1) = u(\tilde{c}_2)$. By [A2] and [A3] $R/\rho > R(\mu/\beta) > R > 1$, which implies $\tilde{c}_2 > c_1$ and hence, $d(c_1) = u(\tilde{c}_2) > u(c_1)$.

(ii) By (i) the incentive constraint (87) does not bind. Let $(1-\pi)\lambda$ and $(1-\pi)\beta\zeta$ be the Lagrange multipliers associated with (86) and (88), respectively. After some rearrangements, the first-order conditions imply:

$$\begin{aligned} \pi[u'(c_1) - \mu/\beta] &= (1-\pi)\lambda d'(c_1) \quad (89) \\ (1+\lambda)Ru'(c_2) &= 1 \quad (90) \\ \lambda + \zeta &= \mu/\beta - 1 \quad (91) \end{aligned}$$

Suppose $\lambda = 0$. Then (89) and (90) simplify to (52) and (53), i.e., the bank contract delivers (c_1^B, c_2^B). By Proposition 6, such a contract violates (86), a contradiction.

(iii) Given $\lambda > 0$, (52) and (89) imply $u'(c_1) > \mu/\beta = u'(c_1^B) > 1$ and so $c_1 < c_1^B < c_1^*$; (53) and (90) imply $Ru'(c_2) < 1 = Ru'(c_2^B)$ and so $c_2 > c_2^B = c_2^*$. ■

Let us compare bank contract (ALT) with the one developed in Section 4, which we will label here as (OPT). Below, we reproduce a more general version of the bank problem, which allows the bank to also set aside cash for evening withdrawals, $z_2^+ \geq 0$. As we argued in Section 4, this option is redundant in the optimal contract, as it is optimal to set $z_2^+ = 0$, but we allow it here to properly compare the two bank contracts. The (generalized) problem of the bank we analyzed in Section 4 is:

$$W(z) \equiv \max_{x,k,\tilde{c}_2,z_2^+} \; z - x - k + \pi u(c_1) + (1-\pi)\tilde{\eta}\left[u(\tilde{c}_2) + c_1 - (\rho/R)\tilde{c}_2\right] + (1-\pi)(1-\tilde{\eta})\left[u(c_2) + \beta z_2^+\right] + \beta W(0) \tag{OPT}$$

subject to

$$\begin{aligned} u(\tilde{c}_2) + c_1 - (\rho/R)\tilde{c}_2 &\geq u(c_2) + z_2^+ && (92) \\ u(c_2) + z_2^+ &\geq u(c_1) && (93) \\ c_1 - (\rho/R)\tilde{c}_2 &\geq 0 && (94) \\ z_2^+ &\geq 0 && (95) \end{aligned}$$

and where

$$\begin{aligned} c_1 &= \frac{(\beta/\mu)x - (1-\pi)(1-\tilde{\eta})\beta z_2^+}{\pi + (1-\pi)\tilde{\eta}} \\ c_2 &= \frac{Rk}{(1-\pi)(1-\tilde{\eta})} \end{aligned}$$

An allocation in either contact (ALT) or (OPT) consists of a morning deposit, afternoon and evening consumption, and money savings. Note that due to the linearity in preferences, the composition of the deposit is not relevant for welfare. Below, we show formally that investors prefer contract (OPT).

Proposition 12 *Consider banking contracts (ALT) and (OPT). Then, (i) an allocation* $\{x + k, c_1, c_2, z_2^+\}$ *that is incentive-feasible in problem (ALT) is also incentive-feasible in problem (OPT); (ii) the converse is not true; and (iii) depositors strictly prefer bank contract (OPT) to (ALT).*

Proof. (i) Given preferences, any effort level is feasible in either bank contract and efforts to obtain cash or capital are perfect substitutes. The marginal cost of effort is the same regardless of total effort level and how that effort is used.

Fix an allocation $\{\bar{x} + \bar{k}, \bar{c}_1, \bar{c}_2, \bar{z}_2^+\}$ that is incentive-feasible in problem (ALT). In both contracts, delivering afternoon consumption $\bar{c}_1$ to impatient depositors carries the same per unit effort penalty 1 and the same return, β/μ. Similarly, the two contracts are equivalent in how $\bar{z}_2^+$ is delivered to all patient depositors (whether directly in the evening or as unspent cash when withdrawing in the afternoon). Note that both problems have the same non-negativity constraints on the unspent cash holdings carried to the following period by patient depositors. The only relevant question then is whether we can deliver $\bar{c}_2$ to *all* patient depositors in problem (OPT) without incurring in more effort than in (ALT). In problem (ALT), there is a cost of 1 unit of effort to obtain R units of evening consumption. In the

(OPT) contract, to deliver evening consumption to depositors without market access, there is a per unit cost of 1 and a benefit of R, same as in (ALT). To deliver evening consumption to depositors with market access, the cost is also 1 in terms of effort, but the benefit is R/ρ; by [A2] and [A3] we have $R/\rho \geq R(\mu/\beta) > R$. Hence, delivering $\bar{c}_2$ to depositors with market access in (OPT) costs less effort than in (ALT) and so, $\{\bar{x}+\bar{k}, \bar{c}_1, \bar{c}_2, \bar{z}_2^+\}$ is feasible in (OPT).

The incentive constraint (87) is the same as (93). So, we are left to check that a feasible allocation $\{\bar{x}+\bar{k}, \bar{c}_1, \bar{c}_2, \bar{z}_2^+\}$ that satisfies (86) also satisfies (92). Setting $c_1 = \bar{c}_1$, $c_2 = \tilde{c}_2 = \bar{c}_2$ and $z_2^+ = \tilde{z}_2^+ = \bar{z}_2^+$, (92) becomes $u(\bar{c}_2) + \bar{c}_1 - (\rho/R)\bar{c}_2 \geq u(\bar{c}_2) + \bar{z}_2^+$. Recall that $\tilde{z}_2^+ = c_1 - (\rho/R)\tilde{c}_2$ and so, $\tilde{z}_2^+ = \bar{c}_1 - (\rho/R)\bar{c}_2$. Hence, the incentive condition is trivially satisfied.

(ii) Showing that the converse is not true is straightforward: bank contract (OPT) allows for c_2 and $\tilde{c}_2$ to be different, which is clearly not feasible in contract (ALT).

(iii) By (i) any incentive-feasible allocation in (ALT) is also incentive feasible in (OPT). And as shown in Proposition 7, $c_2 < \tilde{c}_2$, which is not feasible in contract (ALT). It follows that (OPT) must yield strictly higher welfare than (ALT). ■