ON THE INSTABILITY OF BANKING AND FINANCIAL INTERMEDIATION*

Chao Gu  Cyril Monnet  Ed Nosal
University of Missouri  University of Berne  FRB Atlanta

Randall Wright
University of Wisconsin

April 8, 2019

Abstract

Are financial intermediaries inherently unstable? If so, why? What does this suggest about government intervention? To address these issues we analyze whether model economies with financial intermediation are particularly prone to multiple, cyclic, or stochastic equilibria. Four formalizations are considered: a dynamic version of Diamond-Dybvig incorporating reputational considerations; a model with delegated monitoring as in Diamond; one with bank liabilities serving as payment instruments similar to currency in Lagos-Wright; and one with Rubinstein-Wolinsky intermediaries in a decentralized asset market as in Duffie et al. In each case we find, for different reasons, that financial intermediation engenders instability in a precise sense.

JEL Classification Numbers: D02, E02, E44, G21

Keywords: Banking, Financial Intermediation, Instability, Volatility

*For input we thank George Selgin, Guillaume Rocheteau, Steve Williamson, David Andolfatto, Allen Head, Alberto Teguia and Yu Zhu. Wright acknowledges support from the Ray Zemon Chair in Liquid Assets at the Wisconsin School of Business. The usual disclaimers apply.
Banks, as several banking crisis throughout history have demonstrated, are fragile institutions. This is to a large extent unavoidable and is the direct result of the core functions they perform in the economy. *Finance Market Watch Program @ Re-Define, Banks: How they Work and Why they are Fragile.*

**Introduction**

It is often said that banks, or more generally financial intermediaries, are inherently unstable and prone to volatility. This seems to be based on the notion that financial institutions are “special” compared to, say, producers or middlemen in retail. Keynes (1936), Kindleberger (1978) and Minsky (1992) are names associated with such a position, with Williams (2015) providing a recent perspective (see also Akerlof and Shiller 2009 or Reinhart and Rogoff 2009). Rolnick and Weber (1986) provide evidence of the widespread acceptance of this view when they say: “Historically, even some of the staunchest proponents of laissez-faire have viewed banking as inherently unstable and so requiring government intervention” (as a leading case, Friedman 1960, who defended unfettered markets in virtually all other contexts, advocated bank regulation in his program for monetary stability). As additional evidence, consider the voluminous literature dedicated to the study of bank runs.¹

We share an interest in the questions with which Gorton and Whinton (2002) start their well-known survey: “Why do financial intermediaries exist? What are their roles? Are they inherently unstable? Must the government regulate them?” While there are different ways to proceed, our approach is to build formal models of the institutions and see if they are particularly prone to multiple equilibria or volatile dynamics, including cyclic, chaotic or stochastic outcomes that entail fluctuations even if fundamentals are constant. Central to this approach, by models of

¹For now we discuss bank runs, panics, financial crises, etc. without defining these formally. As Rolnick and Weber (1986) put it, “There is no agreement on a precise definition of inherent instability in banking. However, the conventional view is that it means that general bank panics can occur without economy-wide real shocks.” They add “The usual explanation... involves a local real economic shock that becomes exaggerated by the actions of incompletely informed depositors,” and suggest this is consistent with Friedman and Schwartz’s (1963) view. In terms of models, Chari and Jagannathan (1988) have withdrawals by informed depositors lead to withdrawals by others, while Gu (2011) formalizes this as rational herding. Our approach is different, and avoids fixating only on runs, but we focus squarely on volatility “without economy-wide real shocks.”
intermediation we mean more than models *with* intermediation. It does not suffice to assert, say, that households lend to banks and banks lend to firms but households do not lend to firms – that may be a model *with* banking but not *of* banking.\(^2\)

While there is much work on financial intermediation, there is no generally-accepted, all-purpose model. This is because the institutions perform a myriad of functions that are difficult to capture in a single setup: they serve as middlemen between savers and borrowers or asset sellers and buyers; they screen and monitor investment opportunities on behalf of depositors; they issue liabilities like demand deposits that facilitate third-party transactions; they provide liquidity insurance or maturity transformation; they are safe keepers of cash and other valuables; and they maintain privacy about their assets or their customers. Different approaches are typically used to model these diverse activities, and, in this tradition, we consider several distinct specifications. All of these are constructed using building blocks taken from off-the-shelf models, although the ways in which we combine and apply them are novel, as dictated by the applications at hand.

The first formulation extends Diamond and Dybvig’s (1983) model of liquidity insurance to an infinite horizon, to highlight bankers’ reputation as in Gu et al. (2013a), which is itself based on the model of unsecured credit in Kehoe and Levine (1993). The second features fixed costs of exploiting investment opportunities, similar to Diamond (1984) and Huang (2017). The third, an adaptation of Nosal et al. (2017), puts intermediaries like those in Rubinstein and Wolinsky (1987) into an OTC asset market similar to Duffie et al. (2005). The fourth has bank liabilities serving as payment instruments, similar to currency in Lagos and Wright (2005) and Berentsen et al. (2007), in an environment where bank liabilities are less sus-

\(^2\)The declaration that households lend to banks and banks lend to firms but households do not lend to firms is reminiscent of monetary economics following Clower (1965), who said money buys goods and goods buy money but goods do not buy goods. While once a popular shortcut, it is hard to argue that Clower (cash-in-advance) constraints constitute the last word in monetary theory, and we feel similarly about banking (see Wright 2017 for more on this). Of course it is not necessary for every study to make everything endogenous – e.g., Debreu (1959) made progress using a theory *with* firms and households but not a theory *of* firms and households – but it is crucial, we think, to have financial institutions emerge endogenously when asking if they are unstable as “the direct result of the core functions they perform” (from the epigraph).
ceptible to loss or theft, as in He et al. (2007) and Sanches and Williamson (2010), or less sensitive to information, as modeled by Andolfatto and Martin (2013), Dang et al. (2017) and others mentioned below.

We find in each case that financial intermediation can indeed engender instability: an economy with these institutions has multiple equilibria or volatile dynamics for more parameters than the same economy without them. In some cases there is a unique equilibrium without intermediation and multiple or volatile equilibria with it; in other cases both can have multiple or volatile equilibria but intermediation expands the set of parameters for which this is the case. Further, while the economic logic differs across models, in each case the results are directly related to the raison d’être for intermediation.

As the literature on financial intermediation is vast, we refer to standard references (e.g., Freixas and Rochet 2008; Calomiris and Haber 2014; Vives 2016). We do mention Shleifer and Vishny (2010), which has a similar title and provides additional motivation, even if our methods are quite different, coming mainly from monetary theory as surveyed by Lagos et al. (2017) or Rocheteau and Nosal (2017). In particular, we study infinite-horizon economies because our interest is in economic dynamics, and moreover, finite-horizon models are ill suited for capturing salient aspects of financial activity, including unsecured credit, reputational considerations, and money.\footnote{If there is a terminal date $T$, then at $T$ no one honors obligations to preserve their reputations, nor do they accept currency; so no one does at $T - 1$, and so on – at least without ad hoc devices, e.g., imposing exogenous default penalties or putting money in utility functions.} We also use general equilibrium, in the sense of logically closed systems, without (as much as possible) exogenous assumptions on prices, contracts or behavior. This is because we want to know if instability arises from intermediation \textit{per se} and not extraneous features like noise traders, irrational expectations, etc. To be clear, we impose frictions such as limited commitment, imperfect information or communication, and spatial separation, but those are imposed on the environment, not on prices, contracts or behavior. Thus, we think, what follows are models of financial intermediation, not merely those \textit{with} financial intermediation.
Model 1: Insurance

As mentioned, while our various specifications are mutually distinct, they all make use of standard building blocks from the literature, to make it clear that we are not introducing features that have not been previously deemed relevant. The first setup extends Diamond and Dybvig’s (1983) model of liquidity insurance, or maturity transformation, to an infinite horizon world. As in Gu et al. (2013a), this lets us incorporate reputational considerations à la Kehoe and Levine (1993).

There are some agents that live forever, plus at each date in discrete time there is a [0, 1] continuum of agents that are around only for that period, which is a simple way to have agents care differently about reputations. Each period has two subperiods. The short-lived agents are homogeneous ex ante but face idiosyncratic shocks: any individual is impatient with probability \( \pi \) and patient with probability \( 1 - \pi \), where the impatient (patient) ones derive utility only from consumption in the first (second) subperiod. Given the shock, which is private information, short-lived agents have utility \( u_j(c_j), j = 1, 2 \), where \( c_j \) is the single consumption good in subperiod \( j \), with \( u'_j > 0 \) and \( u''_j < 0 \). \(^4\) Infinite-lived agents have period utility \( v(c) \) for \( c \) in the first or second subperiod, with \( v' > 0, v'' \leq 0 \) and \( v(0) = 0 \).

Short-lived agents have an endowment of 1; infinitely-lived agents have 0. The (completely standard) technology is this: a unit of the good invested at the start of the first subperiod yields \( R > 1 \) units in the second subperiod, or the investment can be terminated at the end of the first subperiod to get back the input. The good can also be stored one-for-one across subperiods. As in any Diamond-Dybvig model, to insure against the shocks, the short-lived agents can form a coalition that resembles, and is interpreted as, a bank. Thus they design a contract \( (c_1, c_2) \) to solve

\[
\begin{align*}
\max_{c_1, c_2} \{ \pi u_1(c_1) + (1 - \pi) u_2(c_2) \} \\
st \quad (1 - \pi) c_2 = (1 - \pi c_1) R \\
\quad c_2 \geq c_1,
\end{align*}
\]

\(^4\)Many applications of Diamond-Dybvig assume \( u_1(\cdot) = u_2(\cdot) \), but not all (e.g., Peck and Shell 2003). The flexibility of the general version is useful for constructing illustrative examples.
where (2) is feasibility and (3) is a truth-telling constraint (if $c_2 < c_1$ patient agents would claim to be impatient, get $c_1$ and store it to the next subperiod). There are also nonnegativity constraints omitted to save space.

This problem is well understood. One result is: assuming $u'_1 (1) > u'_2 (R) R$ and $u'_1 (c) \leq u'_2 (c) R$ at $c = R/(1 - \pi + \pi R)$, we get $1 < c_1^* < c_2^* < R$, so (3) is not binding, and full insurance/efficiency obtains. However, this requires commitment; otherwise, when they learn they are patient and are supposed to make transfers to the impatient, agents will renege. Given our short-lived agents cannot commit, naturally, there emerges a role for long-lived agents as bankers who accept deposits, invest them, and pay off depositors on demand at terms to be determined. Importantly, bankers do not have exogenous commitment ability – it is endogenous and based on reputation. Thus, bankers honor obligations lest they get identified as renegers, whence they are punished to autarky, which is a credible threat because there are many perfect substitutes for any given banker.

In particular, a banker may be tempted to misbehave as in the “cash diversion” models in Biais et al. (2007) or DeMarzo and Fishman (2007): if he misappropriates $d$ deposits, he gets payoff $\lambda d$, where $\lambda$ is not too big, so this is socially inefficient, but he might do it opportunistically. As in Gu et al. (2013a,b), the cost is that he gets caught, and punished, with probability $\mu \leq 1$, where one interpretation is that $\mu$ is the probability one generation of depositors can communicate banker misbehavior to the next generation.\footnote{While $\mu = 1$ is fine, it does not simplify much, and it is known from other applications that $\mu < 1$ can be interesting.} In any case, depositors may set $d < 1$, and invest $1 - d$ on their own, to reduce bank incentives to misbehave, different from most papers that simply assume $d = 1$, but similar to Peck and Setayesh (2019). In addition to $d$, the contract now specifies payouts per deposit contingent on withdrawal time $r_j$, $j = 1, 2$, and the banker’s income $b \in [0, d]$, which he invests for utility $v(b R)$.

Since there is more than one long-lived agent, the short-lived agents can choose any of them to act as banker, and in the spirit of Diamond-Dybvig they make this choice as a coalition. However, we assume they can choose only one, to avoid
determining the optimal number of bankers, because while that is interesting (see Model 2) it would be a distraction here; it can be rationalized by assuming that it is too costly to monitor more than one. Still, since they can choose any one, for reasons often summarized as Bertrand competition, the contract maximizes the expected utility of the depositors. Still, a banker may get positive surplus – rent on his option to act opportunistically – because the contract must give him incentives to not to abscond with the deposits.

The banker’s incentive constraint is

\[ v(b_t R) + \beta V_{t+1} \geq \lambda d_t + \beta (1 - \mu) V_{t+1}, \]  

(4)

where \( \beta \) is his discount factor, \( V_t \) is his equilibrium payoﬀ, and the RHS is the deviation payoﬀ, including \( \lambda d \) for sure and \( V_{t+1} \) iﬀ he is not caught. Note that \( V_{t+1} \) is his valuation next period, facing a new generation of depositors, and hence is taken as given when designing a contract at \( t \). Also note that bankers do not run oﬀ with \( d \) on the equilibrium path, but if one were to, he would get \( d \) and not \( v(Rb) \) – i.e., his income \( b \) is treated the same as any other funds under his control.

This leads to a new contracting problem

\[
\begin{align*}
\max_{d_t, r_{1t}, r_{2t}, b_t} & \{ \pi u_1 (d_t r_{1t} + 1 - d_t) + (1 - \pi) u_2 [d_t r_{2t} + (1 - d_t) R] \} \\
\text{st} \quad (1 - \pi) d_t r_{2t} &= (d_t - b_t - \pi d_t r_{1t}) R \\
\lambda d_t - v(b_t R) &\leq \phi_t,
\end{align*}
\]  

(5)

(6)

(7)

(8)

where (8) rewrites (4) using \( \phi_t \equiv \beta \mu V_{t+1} \). Note that \( \phi_t \) is a bank’s franchise value, capturing the banker’s reputation for trustworthiness. Substituting (6) into (5) to eliminate \( r_2 \), ignoring \( t \) subscripts for now, and taking FOC’s wrt \( (r_1, d, b) \) we get

\[
\begin{align*}
    r_1 & : \quad d \{ \pi [u'_1 (c_1) - Ru'_2 (c_2)] - \eta_1 (1 - \pi + R \pi) \} r_1 = 0 \\
    d & : \quad \{ (r_1 - 1) \pi [u'_1 (c_1) - Ru'_2 (c_2)] + \eta_1 [R - (1 - \pi + R \pi) r_1] - \eta_2 \lambda \} d = 0 \\
    b & : \quad [-u'_2 (c_2) - \eta_1 + \eta_2 v' (b R)] b = 0,
\end{align*}
\]
where \( c_1 = dr_1 + 1 - d \) and \( c_2 = dr_2 + (1 - d) R \), while \( \eta_1 \) and \( \eta_2 \) are multipliers for constraints (7) and (8).

These FOC's yield two critical values, \( \phi^* > 0 \) and \( \hat{\phi} < \phi^* \), delineating three regimes. (i) If \( \phi \geq \phi^* \) then (8) is slack, and \( b = 0 \), since the franchise value keeps the banker honest without \( b > 0 \). In this case there is a continuum of contracts achieving the full-insurance outcome, since depositors can have the bank invest a lot or a little, and in the latter case invest the rest on their own (exactly as in Peck and Setayesh 2019). (ii) If \( \phi \in [\hat{\phi}, \phi^*] \) we must either lower \( d < d^* \) or raise \( b > 0 \) to satisfy (8). While lowering \( d \) from \( d^* \) means less-than-full insurance, this is second-order by the envelope theorem, so we set \( d = \phi/\lambda \) and keep \( b = 0 \). (iii) If \( \phi < \hat{\phi} \), lowering \( d \) further entails too much risk, so we set \( b > 0 \). In case (i), one of the payoff-equivalent contracts has \( r_1 = r_2 \), and (ii)-(iii) the unique contract has \( r_1 = r_2 \); hence wlog we set \( r_1 = r_2 = r \) from now on.

In regime (iii), e.g., \((d, r, b)\) satisfies

\[
\begin{align*}
    b &= d/R [(R - (1 - \pi + R\pi) r] \\
    \phi &= \lambda d - v(bR) \\
    \frac{u'_2(c_2)}{u'_1(c_1) - Ru'_2(c_2)} &= \frac{\pi}{1 - \pi + R\pi} \left[ \frac{(R - 1)(1 - \pi)}{\lambda} v'(bR) - 1 \right]
\end{align*}
\]

with \( c_1 = 1 - d + (d - b) R/(1 - \pi + R\pi) \), \( c_2 = (1 - d) R + (d - b) R/(1 - \pi + R\pi) \).

These and the analogs from regimes (i)-(ii) characterize the contract given \( \phi \), and in particular, one can easily check \( b'(\phi) < 0 \ \forall \phi < \phi^* \), which is important below. This is shown Fig. 1 for the following parameterization:

**Example 1:** Let \( v(b) = Bb \),

\[
\begin{align*}
    u_1(c_1) &= A_1 \frac{(c_1 + \varepsilon)^{1-\sigma_1} - \varepsilon^{1-\sigma_1}}{1 - \sigma_1} \quad \text{and} \quad u_2(c_2) = A_2 \frac{(c_2 + \varepsilon)^{1-\sigma_2} - \varepsilon^{1-\sigma_2}}{1 - \sigma_2},
\end{align*}
\]

where \( B = 0.95 \), \( \sigma_1 = \sigma_2 = 2 \), \( \varepsilon = 0.01 \), \( A_1 = 1 \), \( A_2 = 0.1 \), \( R = 2.1 \), \( \mu = 0.7 \), \( \pi = 0.25 \), \( \lambda = 0.6 \) and \( \beta = 0.99 \).

---

\(^6\)Notice \( \hat{\phi} > 0 \) here (in fact, for the example \( \hat{\phi} = 0.3257 \) and \( \phi^* = 0.600 \)); the case \( \hat{\phi} < 0 \) is less interesting because it never has banking in steady state. In terms of primitives, one can show that \( \hat{\phi} > 0 \text{ iff } \pi [u'_1(1) - Ru'_2(R)] [(R - 1)(1 - \pi) v'(0) - \lambda] > u'_2(R) (1 - \pi + R\pi) \lambda \).
As mentioned, the contract takes $\phi$ as given. To embed this in general equilibrium, use $\phi_t \equiv \beta \mu V_{t+1}$ to write $V_t = v(b_t R) + \beta V_{t+1}$ as a dynamical system,

$$\phi_{t-1} = f(\phi_t) \equiv \beta \mu v[b(\phi_t) R] + \beta \phi_t,$$

(12)

where the function $b(\phi_t)$ comes from the contracting problem. Equilibrium is defined as a nonnegative, bounded path for $\phi_t$ solving (12), from which the other endogenous variables follow using the FOC’s. A stationary equilibrium, which is the same as a steady state here, solves $\phi = f(\phi)$. The nature of steady state depends on whether $\hat{\phi} \leq 0$ or $\hat{\phi} > 0$ (conditions for which are given in fn. 6). Appendix A proves:

**Proposition 1** If $\hat{\phi} \leq 0$ the unique steady state has no banking, $d = 0$. If $\hat{\phi} > 0$ the unique steady state has $\tilde{\phi} \in (0, \hat{\phi})$ and banking, $d > 0$. 

For dynamic equilibria, first note from (12) that $f'(\phi_t)$ has an increasing linear term and a decreasing nonlinear term since (as remarked above) $b'(\phi) < 0$. If the net effect implies $f'(\phi_t) < 0$ over some range the system can exhibit nonmonotone dynamics. For Example 1, Fig. 2a shows $f$ and $f^{-1}$, which cross on the 45° line at $\phi = 0.3215$. In this case the system is monotone and there is a unique equilibrium, which is the steady state, because that is the only bounded path solving (12). To see this is not true in general, consider:

**Example 2:** Same as above except $\sigma_1 = 10$ and $\mu = 1$. 
As Fig. 2b shows, Example 2 implies \( f'(\phi) < -1 \), and so \( f \) and \( f^{-1} \) intersect off the 45\(^\circ\) line, at \((\phi_L, \phi_H)\) and \((\phi_H, \phi_L)\) with \( \phi_H = 0.0696 \) and \( \phi_L = 0.0689 \). As is standard (see Azariadis 1993), this means there is a two-cycle equilibrium where \( \phi_t \) oscillates deterministically between \( \phi_L \) and \( \phi_H \); it also means there are sunspot equilibria where \( \phi_t \) fluctuates randomly between values close to \( \phi_L \) and \( \phi_H \) (see Appendix B). Thus we can get deterministic or stochastic volatility with banking and not without it. That does not mean banking is a bad idea, as it provides insurance to agents who cannot get it otherwise, due to commitment issues. It does mean banking can engender instability.

Figure 2a: Model 1, monotone \( f \)  
Figure 2b: Model 1, nonmonotone \( f \)

Figure 3: Model 1, time series for a two-cycle
The intuition is straightforward: if next period $V_{t+1}$ is high then this period $\phi_t$ is high and we can discipline bankers with low $b_t$; but that makes the current $V_t$ and hence $\phi_{t-1}$ low. This simple logic induces a tendency towards oscillations, but for a cycle the effect has to dominate the linear term in $f(\phi_t)$, which is why parameters matter. Fig. 3 plots time series of $(\phi, d, b, r)$ over the cycle in Example 2. Notice $r$ moves with $\phi$ and $b$ against $\phi$. Whether $d$ moves with or against $\phi$ depends on parameters, but here it is the latter. While the point is not to take this example seriously in a quantitative sense, obviously, it is worth noting that the theory does make qualitative predictions, and does not say that “anything goes.”

![Figure 4: Model 1, two- and three-cycles](image)

Fig. 4 displays the existence of two-cycles in a different way, as fixed points of the second iterate $f^2 = f \circ f$, for another parameterization:

**Example 3:** $B = 1$, $\sigma_1 = 14$, $\sigma_2 = 1.5$, $\varepsilon = 0.01$, $A_1 = 1$, $A_2 = 0.075$, $R = 2.2$, $\mu = 1$, $\pi = 0.28$, $\lambda = 0.75$ and $\beta = 0.76$.

Notice $f^2$ has three fixed points, $\bar{\phi}$, plus $\phi_L$ and $\phi_H$ from the two-cycle. Also shown is $f^3$, which has seven fixed points, $\bar{\phi}$ plus a pair of three-cycles. Standard results (again see Azariadis 1993) say the existence of three-cycles implies the existence of
cycles for any $n$, plus chaos, which is basically a cycle with $n = \infty$.

To summarize, banking can introduce many equilibria, including deterministic, stochastic and chaotic dynamics, directly attributable to the idea that banks depend on trustworthiness, and at least to some extent that is a self-fulfilling prophecy. We have more to say about this after presenting some other models.

Model 2: Delegated Investment with Fixed Costs

The next formulation has intermediation originating from economies of scale, based on Diamond (1984) and Huang (2017) (see also Leland and Pyle 1976 or Boyd and Prescott 1986 on the bigger picture). Time is discrete and continues forever as in Model 1, but here all agents are infinitely lived. Also, they are now spatially separated – say, across a large number of islands – and randomly relocated at the end of each period, following a literature on banking including Champ et al. (1996), Bencivenga and Smith (1991), Smith (2002) and Bhattacharya et al. (2005). Economies of scale are captured as follows: agents must pay a fixed cost $\kappa$, in terms of goods, to locate/evaluate/monitor investment projects, after which any project returns $R$ per unit invested.

Period utility is $u(x) - c(d)$, where $x$ is consumption and $d$ investment – say, output produced one-for-one with labor – where $u', c' > 0$ and $c'' \geq 0 > u''$. Also, $u'(0)R > c'(0)$, so that agents invest if $\kappa = 0$. If $\kappa > 0$ the payoff is

$$W_1 = \max_{x,d} \{u(x) - c(d)\} \text{ st } x = Rd - \kappa,$$

from investing on one’s own (omitting nonnegativity constraints as usual). Suppose $\kappa$ is too high to support this, so $W_1 < 0$, and denote the autarky payoff by $W_0 = 0$. Now consider agents forming a coalition where some, that we call depositors, delegate their investment to others, that we call bankers, to share the fixed cost.\footnote{The main function of random relocation here is to let us avoid long-term contracting considerations, which are interesting but complicated (e.g., in Gu et al. 2013a, bankers’ rewards can be backloaded over multi-period contracts). Elsewhere in the paper we avoid those issues using short-lived agents, but here we want all agents to be long-lived, so that ex ante anyone can potentially be a banker. In any case, it is important to emphasize that these are not restrictions on contracting per se, but assumptions on the environment that impinge on the contract. Does it matter? Yes, because without making them explicit one cannot know, in general, how these assumptions impinge on all endogenous variables.}
As in many models with nonconvexities, a coalition uses a lottery to choose a subset of members to act as bankers.\textsuperscript{8} Thus, $\omega_t$ is the probability of being a banker, and the measure of bankers if the island population is normalized to 1. As in Model 1, bankers have the option to misbehave, with $\lambda$ and $\mu$ playing similar roles. The relevant incentive condition is therefore

$$\beta V_{t+1} \geq \frac{\lambda (1-\omega_t) x_t}{\omega_t} + (1-\mu) \beta V_{t+1},$$

(14)

where the RHS is the deviation payoff, given each depositor is promised $x_t$ and each banker controls $(1-\omega_t) x_t / \omega_t$ of the resources.\textsuperscript{9} The trade-off, as in Huang (2017), is that having fewer banks saves on fixed costs but raises their temptation to misbehave, because they must be large, unless we lower total deposits.

The contract maximizes the payoff to the representative agent on the island

$$W(\phi) = \max_{\omega,X,D,x,d} \left\{ \omega \left[ u(X) - c(D) \right] + (1-\omega) \left[ u(x) - c(d) \right] \right\}$$

(15)

subject to

$$\omega X + (1-\omega) x = R \left[ \omega D + (1-\omega) d \right] - \kappa \omega$$

(16)

$$u(x) - c(d) \geq 0$$

(17)

$$\frac{1-\omega}{\omega} x \leq \phi,$$

(18)

where $\phi_t \equiv \mu \beta V_{t+1} / \lambda$, while $(X, D)$ and $(x, d)$ are the consumption/investment allocations of bankers and depositors. Here (16) is the resource constraint, (17) is the incentive constraint for depositors, and (18), which is the same as (14), is the incentive constraint for bankers.

Substituting (16) into (15) to eliminate $X$, and letting $\eta$ and $\gamma$ be multipliers, we get the FOC’s:

$$D : \ u'(X) R - c'(D) = 0$$

$$x : \ (1-\omega) [u'(X) R - c'(d)] - \eta c'(d) = 0$$

$$\omega : \ \omega \left\{ u(X) - c(D) - [u(x) - c(d)] + u'(X) \frac{x - Rd}{\omega} + \gamma \frac{x}{\omega^2} \right\} = 0$$

\textsuperscript{8}This is similar to, e.g., Rogerson’s (1988) indivisible-labor model, except unlike his, our agents cannot commit, so our contracts must be incentive compatible before and after the lottery.

\textsuperscript{9}Here we assume that a deviating banker receives $u(X) - c(D)$ in addition to $\lambda[(1-\omega_t)x_t/\omega_t]$. 

13
One can check $W'(\phi) \geq 0$. Moreover, $W(0) = 0$, so we get no banking at $\phi = 0$. In the limit as $\phi \to \infty$ we get $\omega \to 0$, which means very few banks but they are huge. Also as $\phi \to \infty$ we get $W(\phi) \to W^* \equiv \max_{x,d} [u(x) - c(d)]$ st $x = Rd$, which totally dissipates the fixed cost (i.e., delivers the same payoff as $\kappa = 0$).

Fig. 5 shows the contract given $\phi$ for the following parameterization:

**Example 4:** Let

$$u(x) = A(x + \varepsilon)^{1-\sigma} - \varepsilon^{1-\sigma}$$

and $c(d) = Bd$,

with $A = \varepsilon = 0.001$, $\sigma = 2$, $B = 0.1$, $\kappa = 230$, $R = 1.2$, $\beta = 0.76$, $\mu = 0.95$ and $\lambda = 9$.

Notice that there is a cutoff $\bar{\phi}$, which happens to be $\bar{\phi} = 0.0182$ in this example, and banking is viable iff $\phi \geq \bar{\phi}$.

![Figure 5: Model 2, bank contract vs $\phi$](image)

To embed this in equilibrium we use $V_t = W(\phi_t) + \beta V_{t+1}$ and emulate the methods from Model 1 to get

$$\phi_t = f(\phi_{t+1}) \equiv \frac{\beta \mu}{\lambda} W(\phi_{t+1}) + \beta \phi_{t+1}. \quad (19)$$

Equilibrium is a bounded, nonnegative solution to (19). Notice $f(\phi) = \beta \phi$ for $\phi \leq \bar{\phi}$, and $f(\phi) < \phi$ for big $\phi$ due to the fact that $W \leq W^*$.\(^10\) Thus we have:

\(^{10}\)For example, if $\phi = \mu W^*/\lambda$, then (19) necessarily implies that $\phi < f(\phi)$. 14
Proposition 2 There is always a steady state with $\phi = 0$, without banking. There can be steady states with banking, generically an even number that alternate between stable and unstable.

Fig. 6 shows Example 4 has three steady states, the ubiquitous $\phi = 0$, plus two with banking $\phi_2 > \phi_1 > 0$. This is different from Model 1, which has a unique steady state $\bar{\phi}$, and has nonstationary equilibria iff $f'(\bar{\phi}) < -1$. Now $f'(\phi) > 0$, so deterministic cycles are impossible, but if there are multiple steady states we can use a different approach to construct sunspot equilibria around the stable ones. Appendix B shows there are equilibria where $\phi$ fluctuates between $\phi_A$ and $\phi_B$ for any $\phi_A \in (\phi_0, \phi_1)$ and $\phi_B \in (\phi_1, \phi_2)$. In particular, $\phi_A < \bar{\phi}$ means we switch stochastically between $d_t > 0$ and $d_t = 0$—i.e., random episodes of crises, where deposits dry up and banking shuts down due to fundamentally irrelevant events (sunspots). This is different from Model 1, where $d_t$ can fluctuate, but only with $d_t > 0 \forall t$.

---

11 To give credit where credit is due, in Model 2 we use the method in Azaridais (1981), while in Model 1 we use the method in Azaridais and Guesnerie (1986).
While Models 1 and 2 are different, in terms of economics and mathematics, Appendix C presents an environment that integrates elements of both. It has two agents on each island, one that is infinitely lived and one that is only around for one period, who bargain over the terms of the contract (having just two is natural for bargaining, but we also considered many depositors and one banker, with multilateral bargaining, and got similar results). There are gains to delegating investment due to $\kappa > 0$, as in Model 2, but now only long-lived agents can act as bankers, as in Model 1. Letting $\theta$ denote bankers’ bargaining power, we get a dynamical system $\phi_t = f(\phi_{t+1})$ that can be nonmonotone for $\theta < 1$. For this system Appendix C shows we can have multiple steady states, with $f'(\phi) > 0$ around the stable ones and hence sunspot equilibria as in the benchmark Model 2, as well as $f'(\phi) < -1$ around the unstable steady states and hence cycles and sunspots as in Model 1.

The reason $f(\phi)$ is decreasing in Appendix C is the following well-known (see Kalai 1977) feature of Nash bargaining: agents with bargaining power $\theta < 1$ can get a smaller surplus when the bargaining set expands. Here this is manifest in bankers’ surplus falling with $\phi$, similar to $b'(\phi) < 0$ in Model 1. That does not happen in the baseline Model 2, where agents in the coalition are ex ante identical and hence treated symmetrically, which means they all get a bigger surplus when $\phi$ increases. Details aside, we conclude there are distinct ways to formalize how banking might engender instability though reputation/trust, but they can be integrated.¹²

Model 3: Asset Market Intermediation

Banks are not the only interesting financial intermediaries. Work following Duffie et al. (2005) studies financial markets using search theory, where agents may trade assets with each other, or with middlemen/dealers that buy from those with low valuation and sell to those with high valuation. We pursue this with a few changes

¹²Model 1 can actually be interpreted as bargaining where banks have $\theta = 0$, since that is basically the same as Bertrand competition. Hence, one may conjecture that similar dynamics emerge in Model 1 if we allow $\theta \in (0,1)$, but then it becomes intractable, while the setup in Appendix C does not. Another model that gets nonmonotone dynamics from $\theta \in (0,1)$ can be found in Gu et al. (2013b, Appendix B), but that has simple borrowing and lending without banks. Still, we put our model with $\theta \in (0,1)$ in the Appendix because it is not the first formulation where nonmonotone dynamics emerge from generalized Nash bargaining.
in their environment. In particular, most papers following Duffie et al. (2005) give middlemen continuous access to a frictionless interdealer market (with some exceptions, e.g. Weill 2007, but they do not study the issues analyzed below). Hence, their intermediaries do hold assets in inventory. Our middlemen are more like those in Rubinstein and Wolinsky (1987), who buy from producers and hold inventories until they randomly sell to consumers. However, here they trade assets and not goods, which matter a lot as argued by Nosal et al. (2017), on which this section is based, even if we amend their setup in many ways. These include: adding heterogeneous projects; modifying the market composition conditions (see fn. 13); and switching from continuous to discrete time, which generates a few new results.

There are large numbers of three risk-neutral types, $B$, $S$ and $M$, for asset buyers, sellers and middlemen. Type $M$ agents stay in the market forever, while type $S$ and $B$ stay for one period, although we also studied alternatives, like letting everyone stay forever, with similar results. Upon exit $S$ and $B$ are replaced by ‘clones’ to maintain stationarity (a device borrowed from Burdett and Coles 1997). Type $B$ agents, sometimes called end users, want to acquire an asset – let’s call it capital – to implement a project for profit $\pi > 0$, where $\pi$ is observable when agents meet, but random across end users with CDF $F(\pi)$. The originators of capital, type $S$, if they enter the market each bring 1 indivisible unit; those that stay out put their capital to alternative uses, defining their opportunity cost of participation, denoted $\kappa_s > 0$, which for simplicity is the same for all $S$.

Type $M$ agents, who are always in the market, can acquire capital from $S$, but as usual in these models their inventories are restricted to $k \in \{0, 1\}$ (with exceptions, e.g. Lagos and Rocheteau 2009, but they do not study the issues analyzed below). Let $n_t$ be the measure of type $M$ at $t$ with $k = 1$. Capital held by $M$ depreciates by disappearing each period with probability $\delta \geq 0$, but while he holds it $M$ gets a return $\rho > 0$. His crucial choice is then, if he has $k = 1$ and meets $B$, should he trade or keep $k$ for himself? This depends on fundamentals, of course, including the end user’s $\pi$, but it can also depend on equilibrium considerations.
Market composition is determined as follows: the measures \( n_m \) and \( n_b \) of types \( M \) and \( B \) are fixed, but entry by \( S \) makes \( n_s \) endogenous.\(^{13}\) Given this, the meeting technology is standard: each period everyone in the market contacts someone with probability \( \alpha \), and each contact is a random draw from the participants. In particular, if \( N_t \) is the total measure of participants then types \( M \) and \( S \) both meet type \( B \) with probability \( \alpha n_b/N_t \), so \( M \) has no advantage over \( S \) that regard – different from the original Rubinstein-Wolinsky setup, it is but fine for our purposes. When \( B \) and \( S \) meet they trade for sure since this is \( S \)’s only chance and cost \( \kappa_s \) is sunk. Similarly, when \( S \) meets \( M \) with \( k = 0 \) they trade for sure. When \( M \) with \( k = 1 \) meets \( B \), as mentioned, they may or may not trade.

As regards prices, if type \( j \) gives \( i \) capital the latter pays \( p_{ij} \) (in terms of transferable utility), determined by bargaining. Thus, if \( \Sigma_{ij} \) is the total surplus available when \( i \) and \( j \) meet, as long as \( \Sigma_{ij} > 0 \) they trade, and type \( i \)’s surplus is \( \theta_{ij} \Sigma_{ij} \), where \( \theta_{ij} \in [0,1] \) is his bargaining power. To flesh this out, let \( V_{s,t} \) and \( V_{b,t} \) be value functions for types \( S \) and \( B \); let \( V_{k,t} \) be the value function for \( M \) when he has \( k \in \{0,1\} \); and let \( \Delta_t = V_{1,t} - V_{0,t} \) be \( M \)’s gain from having inventory. Then

\[
\Sigma_{bs,t} = \pi, \Sigma_{ms,t} = (1 - \delta) \beta \Delta_{t+1}, \Sigma_{bm,t} = \pi - (1 - \delta) \beta \Delta_{t+1},
\]

where \( \beta \in (0,1) \) is \( M \)’s discount factor. Note there are no continuation values or threat points for \( S \) and \( B \), as they are in the market for just one period, but while that simplifies some algebra it is not otherwise important. Bargaining now yields

\[
p_{bs,t} = \theta_{sb} \pi, \quad p_{ms,t} = \theta_{sm} (1 - \delta) \beta \Delta_{t+1}, \quad p_{bm,t} = \theta_{mb} \pi + \theta_{mb} (1 - \delta) \beta \Delta_{t+1}.
\]  \((20)\)

When \( M \) with \( k = 1 \) and \( B \) with project \( \pi \) meet, they trade with probability \( \tau_t = \tau(\pi, R_t) \), where

\[
\tau(\pi, R) = \begin{cases} 
0 & \text{if } \pi < R \\
[0,1] & \text{if } \pi = R \\
1 & \text{if } \pi > R 
\end{cases}
\]  \((21)\)

\(^{13}\)Entry by \( S \) is nice because it lets us compare economies with the same entry conditions with and without middlemen. Still, results for entry by \( M \) are given in Appendix D; entry by \( B \) is less interesting and hence omitted. Arguably these alternatives are all better than Nosal et al. (2017), where agents choose to be either type \( M \) or \( S \). That is awkward because, e.g., in cyclic equilibria they switch back and forth over time between being \( M \) and \( S \). Here no one switches, but participation by a type can vary as in conventional search theory (e.g., Pissarides 2000).
and \( R_t \equiv (1 - \delta) \beta \Delta_{t+1} \) is the reservation value of a project making \( \Sigma_{mb} = 0 \). Hence, the market payoff for \( B \) with project \( \pi \) is

\[
V_{b,t}(\pi) = \frac{\alpha n_{s,t}}{N_t} \theta_{bs} \pi + \frac{\alpha t}{N_t} \tau(\pi, R_t) \theta_{bm} [\pi - (1 - \delta) \beta \Delta_{t+1}].
\]  

(22)

The first term on the RHS is the probability \( B \) meets \( S \), times his share of the surplus; the second is the probability he meets \( M \) with \( k = 1 \), times the probability they trade, times his share of the surplus; and note prices do not appear since they were eliminated using (20). Similarly, the market payoff for \( S \) is

\[
V_{s,t} = \frac{\alpha n_b}{N_t} \theta_{sb} \int_0^\infty \pi dF(\pi) + \frac{\alpha (n_m - n_t)}{N_t} \theta_{sm} (1 - \delta) \beta \Delta_{t+1}.
\]  

(23)

The payoff for \( M \) depends on inventory. Using \( R_t = (1 - \delta) \beta \Delta_{t+1} \), we have

\[
V_{0,t} = \frac{\alpha n_{s,t}}{N_t} \theta_{ms} R_t + \beta V_{0,t+1}
\]  

(24)

\[
V_{1,t} = \rho + \frac{\alpha n_b}{N_t} \theta_{mb} \int_{R_t}^\infty (\pi - R_t) dF(\pi) + (1 - \delta) \beta V_{1,t+1} + \delta V_{0,t+1}.
\]  

(25)

Subtracting and simplifying with integration by parts, we get

\[
R_{t-1} = (1 - \delta) \beta \left\{ \rho + R_t + \frac{\alpha n_b \theta_{mb}}{N_t} \int_{R_t}^\infty [1 - F(\pi)] d\pi - \frac{\alpha n_{s,t} \theta_{ms}}{N_t} R_t \right\}.
\]  

(26)

giving the evolution of \( R \) over time. The evolution of inventories held by \( M \) is

\[
n_{t+1} = n_t (1 - \delta) \left[ 1 - \frac{\alpha n_b}{N_t} \mathbb{E}_T(\pi, R) \right] + \frac{(n_m - n_t) \alpha n_{s,t} (1 - \delta)}{N_t}.
\]  

(27)

where \( \mathbb{E}_T(\pi, R) \) is the unconditional probability that \( M \) and \( B \) trade.\(^{14}\) The first term on the RHS is current \( n \) times the probably a unit of \( k \) does not depreciate or get traded; the second is current \( n_m - n \) times the probability \( M \) acquires \( k \) and it does not depreciate.

We can eliminate \( N_t \) in (26)-(27) using \( S \)'s entry condition, \( V_{s,t} = \kappa_s \), which reduces to

\[
n_{s,t} = \frac{\alpha n_b \theta_{sb} \mathbb{E}_T + \alpha (n_m - n_t) \theta_{sm} R_t}{\kappa_s} - n_b - n_m.
\]  

(28)

\(^{14}\)It is also equal to the probability that \( \pi > R \).
What’s left is a two-dimensional dynamical system that is compactly written as

\[
\begin{bmatrix}
  n_{t+1} \\
  R_{t-1}
\end{bmatrix} = \begin{bmatrix}
  f(n_t, R_t) \\
  g(n_t, R_t)
\end{bmatrix}.
\]  

(29)

Given an initial \( n_0 \), an equilibrium is a nonnegative, bounded path for \((n_t, R_t)\) solving (29).

With no intermediaries, \( n_m = 0 \), as both \( S \) and \( B \) are one-period lived, the equilibrium is static and it is easy to check that it is unique. With \( n_m > 0 \), first note that the locus of points satisfying \( n = f(n, R) \), called the \( n \)-curve, and the locus satisfying \( R = g(n, R) \), called the \( R \)-curve, both slope up in \((n, R)\) space. Then, to develop some intuition, consider the special case where \( \pi = \bar{\pi} \) is constant. As shown in Fig. 7, for this case there are three possible regimes: (i) \( R < \bar{\pi} \), so \( M \) with \( k = 1 \) and \( B \) trade with probability \( \tau = 1 \); (ii) \( R > \bar{\pi} \), so they trade with probability \( \tau = 0 \); and (iii) \( R = \bar{\pi} \), so they trade with probability \( \tau \in (0, 1) \).

Appendix A proves:

**Proposition 3** With \( \pi = \bar{\pi} \) there exist \( \hat{\rho} > 0 \) and \( \check{\rho} > \hat{\rho} \) such that: (i) if \( \rho \in [0, \hat{\rho}) \) there is a unique steady state and it has \( R < \bar{\pi} \); (ii) if \( \rho \in (\hat{\rho}, \infty) \) there is a unique steady state and it has \( R > \bar{\pi} \); (iii) if \( \rho \in (\check{\rho}, \hat{\rho}) \) there are three steady states, \( R < \bar{\pi} \), \( R > \bar{\pi} \), and \( R = \bar{\pi} \).

---

\( ^{15} \)A distinction between this model and others in the paper is that this system is two dimensional: \( R \) is a jump variable, like \( \phi \) in the previous sections, while \( n \) is a (predetermined) state variable, so transitions are nontrivial. Interestingly, the version of Model 3 in Appendix D, with entry by \( M \) instead of \( S \), is different: there one can solve a univariate system \( R_{t-1} = G(R_t) \) to get the path for \( R_t \), after which \( n_t, N_t \) etc. follow from simple conditions. Intuitively, with entry by \( M \) (entry by \( S \)) the model is (is not) block recursive, as discussed in Shi (2009).
For several reasons we prefer a nondegenerate $F(\pi)$.\textsuperscript{16} So, consider a smooth mean-preserving spread of the degenerate case:

**Example 6:** Let

$$F(\pi) = \begin{cases} 
\frac{\pi_1}{\pi_0} & \text{if } 0 \leq \pi \leq \pi_0 \\
\pi_1 + (\pi_3 - \pi_1)(\pi - \pi_0)/(\pi_2 - \pi_0) & \text{if } \pi_0 < \pi \leq \pi_2 \\
\pi_2 + (1 - \pi_3)(\pi - \pi_2)/(\pi_4 - \pi_2) & \text{if } \pi_2 < \pi \leq \pi_4
\end{cases}$$

(30)

with $\pi_0 = 0.99$, $\pi_1 = 0.05$, $\pi_2 = 1.01$, $\pi_3 = 0.95$ and $\pi_4 = 2$. Also, let $\alpha = 1$, $\kappa_s = 0.1$, $n_b = 0.05$, $n_m = 0.5$, $\theta_{sm} = 0.5$, $\theta_{sb} = 1$, $\theta_{mb} = 0.7$, $\beta = 1/1.04$, $\delta = 0.008$ and $\rho = 0.2$.

\textsuperscript{16}For the nondegenerate $F(\pi)$ studied below, the flat portion of the $n$-curve in Fig. 7 is eliminated. Related to this, in any steady state $M$ and $B$ are indifferent to trade only in the rare event $\pi = R$, in contrast to the mixed-strategy equilibrium in the degenerate case, where they are always indifferent. Moreover, with the nondegenerate $F(\pi)$, if $R$ varies across pure-strategy steady states intermediation activity does too, but not necessarily to the extreme extent of the degenerate case, where it is either $\tau = 1$ or $\tau = 0$. Similarly, for real-time dynamics, cycles with nondegenerate $F(\pi)$ have fluctuations in intermediation activity but not necessarily between $\tau = 0$ and $\tau = 1$. Now it may be interesting to have $\tau$ fluctuate between 0 and or 1, which can happen with a nondegenerate $F(\pi)$, but so can less extreme cycles.
This is shown in Fig. 8, which is similar to Fig. 7, but the slope of the $n$-curve is always positive. Hence the results are similar to Proposition 3, including multiplicity $\forall \rho \in (\hat{\rho}, \bar{\rho})$.

Here is the intuition. First suppose $R$ is low, so $M$ trades $k$ to $B$ with a high probability. Then the probability $M$ has $k = 0$ is high, which is good for sellers, so $n_s$ is high. That makes it easy for $M$ to get $k$ and rationalizes his low $R$. Now suppose $R$ is high, so $M$ trades $k$ to $B$ with a low probability. Then the probability $M$ has $k = 0$ is low, which is bad for sellers, so $n_s$ is low. That makes it hard for $M$ to get $k$ and rationalizes his high $R$. Heuristically, this requires multiple intermediaries acting independently – if they could somehow collude (or if there were a monopoly intermediary comprised of many agents, who match like the individual $M$’s in the baseline model, to preserve the nature of the frictions), they could internalize the impact of their $\tau$ decision on entry by $S$. This is different from other models in the paper, where the main results are independent of the number of banks, except to the degree that this may affect the size of their surplus.

Moreover, market liquidity – i.e., the ease with which agents can buy and sell $k$ – is high (low) if $R$ is low (high). Multiplicity means liquidity is not pinned down.
by fundamentals, a recurring theme in monetary theory (e.g., Kiyotaki and Wright 1989), but the intuition here is different. In monetary economies, whether an agent accepts something as medium of exchange depends on what others accept. Here, whether type $M$ agents trade away $k$ depends on $n_s$, and $n_s$ depends on whether type $M$ trade away $k$, which is not the same idea. In particular, our result requires endogenous market composition, which is not true in monetary models.

![Figure 9: Time series for a two-period cycle in Model 3 with entry by $S$](image)

Now consider a two-cycle oscillating between a liquid regime with low $R$ and an illiquid regime with high $R$. That is, we seek $(R^L, n^L)$ and $(n^I, R^I)$ solving

\[
\begin{bmatrix}
  n^I \\
  R^I
\end{bmatrix} = \begin{bmatrix}
  f(n^L, R^L) \\
  g(n^L, R^L)
\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}
  n^L \\
  R^L
\end{bmatrix} = \begin{bmatrix}
  f(n^I, R^I) \\
  g(n^I, R^I)
\end{bmatrix}.
\] (31)

One can verify a solution is $(R^L, n^L) = (0.9862, 0.4504)$ and $(R^H, n^H) = (1.0103, 0.4312)$. Hence, we have real-time dynamics with excess volatility in liquidity, trade volume, prices and quantities. Fig. 9 shows the times series. In the liquid regime: $R$ is low, making $M$ more inclined to trade with $B$; $n$ is high, because $M$ and $B$ traded less last period; and $n_s$ is low, because low $R$ and high $n$ discourage entry by $S$. The
illiquid regime has the opposite properties. Whether output $y$ is higher or lower in the liquid regime depends on parameters.

We do not claim that actual data are best explained by a two-cycle; we do suggest this that if such a simple model can deliver equilibria where liquidity, prices, quantities and other endogenous variables vary over time, as self-fulfilling prophecies, it lends credence to the notion that intermediated asset markets in the real world might be prone to similar instability.\footnote{Prices are also shown in Fig. 9 (averaged over $\pi$ when $B$ trades). The price $B$ pays $S$ is constant over time, as it depends only on fundamentals, but the price $M$ pays $S$ or $B$ pays $M$ moves with $R$. The spread can go either way, but here it moves against $R$. This is all broadly consistent with the data discussed in Comerton-Forde et al. (2010), and other stylized facts (e.g., inventories are volatile than output). While this is obviously not a calibration, the finding that it is qualitatively consistent with several observations may lend further credence to the story.}

**Model 4: Safety and Secrecy**

An important role of banks is the issuance of liabilities that facilitate third-party transactions. Indeed, as conventional wisdom has it, that is virtually their defining characteristic: “banks are distinguished from other kinds of financial intermediaries by the readily transferable or ‘spendable’ nature of their IOUs, which allows those IOUs to serve as a means of exchange, that is, money... Commercial bank money today consists mainly of deposit balances that can be transferred either by means of paper orders known as checks or electronically using plastic ‘debit’ cards” (Selgin 2018). We pursue this idea in a model with an explicit need for payment instruments, building on the New Monetarist framework recently surveyed by Lagos et al. (2017) and Rocheteau and Nosal (2017), where we introduce banks in two related ways.

The first says that bank liabilities are safe relative to other assets in the sense that they are less susceptible to theft or loss (this is similar to He et al. 2007; see Sanches and Williamson 2010). Traveler’s checks are a case in point but not the only one – e.g., it is obviously worse when one has their cash lost or stolen than when a checkbook or debit card is lost or stolen. Similarly, if merchandise turns out to be fraudulent or defective, which is another version to theft, it is typically easier to stop or reverse payments when one pays by check or credit card rather than with
cash. The second version builds on the idea that payment instruments originating with banks can be, as Dang et al. (2017) put it, informationally insensitive when these institutions act as secret keepers (an earlier exposition of this is Gorton and Pennachi 1990; Andolfatto and Martin 2013, Andolfatto et al. 2014 and Loberto 2017 are versions that go deeper into the microfoundations of the transaction process, similar to the way we model it).

While there are different approaches to modeling media of exchange, we use one based on Lagos and Wright (2005), because it is convenient for this application. In that framework, in each period of discrete time, two markets convene sequentially: a decentralized market, or DM, with frictions detailed below; and a frictionless centralized market, or CM, that serves mainly to simplify the analysis. There are two types of infinitely-lived agents, a measure 1 of buyers and a measure $n$ of sellers, whose roles differ in the DM. However, they are similar in the CM, where all agents trade a numeraire consumption good $x$ and labor $\ell$ for utility $U(x) - \ell$, with $U'(x) > 0 > U''(x)$, as well as assets. These assets are interpreted as the standard ‘trees’ in Lucas’s (1978) asset-pricing theory, giving off ‘fruit’ as a dividend $\rho$, denominated in CM numeraire, where as usual fiat currency is the special case $\rho = 0$.

In the DM, where agents meet bilaterally, sellers can provide a good $q$ (different from $x$) that buyers want. Let $\alpha$ be the probability a buyer meets a seller, and $\alpha/n$ the probability a seller meets a buyer. In any meeting, if a seller produces for a buyer the former incurs cost $c(q)$ and the latter gets utility $u(q)$, where $c(0) = u(0) = 0$, $c'(q), u'(q) > 0$ and $c''(q) \geq 0 > u''(q)$. Credit is not viable because goods are nonstorable, there is limited commitment, and the DM trading process is anonymous. Hence, sellers only produce if they get assets in exchange. The terms

---

18Safety was a critical feature of banks historically. Consider the British goldsmiths: “At first [they] accepted deposits merely for safe keeping; but early in the 17th century their deposit receipts were circulating in place of money” (Encyclopedia Britannica, quoted in He et al. 2005). Also, “In the 17th century, notes, orders, and bills (collectively called demandable debt) acted as media of exchange that spared the costs of moving, protecting and assaying specie” (Quinn 1997). Safety was also crucial for earlier bankers, e.g. the Templars (Sanello 2003).

19For other ways of modeling bank liabilities (notes) as media of exchange, that are similar in spirit see, e.g., Cavalcanti and Wallace (1999a,b) or Cavalcanti et al. (1999).
of trade are given by a generic bargaining solution: for a buyer to get \( q \), he must give the sellers assets worth \( v(q) \) in CM numeraire.\(^{20}\)

Assets can be held in forms that differ in safety and liquidity, where safety is measured by the probability of being stolen, and liquidity is measured by whether it is accepted as means of payment in the DM.\(^{21}\) Let \( a = (a_1, a_2) \) be a buyer’s portfolio: \( a_1 \) denotes assets held in a safe but illiquid form, i.e., it cannot be stolen but it cannot be used in the DM either; and \( a_2 \) denotes assets held in a less safe but liquid form, i.e., its probability of being stolen is positive, denoted by \( \delta \), but it can be used as means of payment in the DM.

The ex-dividend price of the asset in terms of numeraire is \( \phi \) independent of what one does with it – i.e., whether one holds it as \( a_1 \) or \( a_2 \). Then a buyer’s state variable at the start of the CM is his wealth \( A = (\phi + \rho) \sum_{j=1}^{2} a_j \) and \( W(A) \) is his value function. However, the DM value function \( V(a) \) depends on his portfolio, not just its value. Thus a buyer’s CM problem at \( t \) is

\[
W_t(A_t) = \max_{x_t, \ell_t, \hat{a}_t} \{ U(x_t) - \ell_t + \beta V_t(\hat{a}_t) \} \text{ st } x_t = A_t + \ell_t - \phi_t \sum_{j=1}^{2} \hat{a}_{j,t}
\]

where \( \hat{a} = (\hat{a}_1, \hat{a}_2) \) is a new portfolio and the CM real wage is 1 because we assume \( x \) is produced one-for-one using \( \ell \), merely to reduce notation. A seller’s problem (not shown) is similar except, unlike buyers, they have no need for liquidity, so they only hold assets that are priced fundamentally (see below). Assuming an interior solution for buyers, we immediately have several standard results: (i) \( x_t = x^* \), where \( U'(x^*) = 1 \); (ii) \( W_t'(A_t) = 1 \), so \( W_t(A_t) \) is linear in wealth; and (iii) \( \hat{a}_t \) is independent of \( a_t \), so all buyers exit the CM with the same portfolio, satisfying the FOC’s

\[
\beta \partial V_{t+1}/\partial \hat{a}_{j,t} \leq \phi_t, = 0 \text{ if } \hat{a}_{j,t} > 0.
\]

While there is a discount factor \( \beta \in (0, 1) \) between the CM and the next DM, wlog agents do not discount between the DM and CM. Thus, a buyer’s value function

---

\(^{20}\)A simple example is Kalai’s (1977) bargaining solution, \( v(q) = \theta c(q) + (1 - \theta) u(q) \). Different from Model 3, Kalai is not the same as Nash bargaining here because utility is nonlinear, and there are reasons to prefer the latter in models of liquidity (Aruoba et al. 2007).

\(^{21}\)Any assets that are stolen (or lost) return to the system next period, either because the thieves (or finders) bring them to the CM, or because a new endowment appears, to maintain stationarity.
is

\[ V_{t+1}(\hat{a}_t) = (1 - \delta) \left\{ W_{t+1}(\hat{A}_{t+1}) + \alpha [u(q_{t+1}) - v(q_{t+1})] \right\} + \delta W_{t+1}(\phi_{t+1} + \rho) \hat{a}_{1,t} \]

where \( \hat{A}_{t+1} \) is the wealth implied by \( \hat{a}_t \), and \( q_{t+1} \) is the quantity transacted when paying with \( \hat{a}_{2,t} \). The buyer’s surplus in the DM transaction is simply \( u(q) - v(q) \), because of the result that \( W(\cdot) \) is linear.

Equilibrium is described by the Euler equations, which we get from inserting the derivatives of the DM value function into the FOC’s from the CM. First, to ease notation let \( \lambda(q) = u'(q)/v'(q) - 1 \) be the liquidity premium (i.e., the Lagrange multiplier on the constraint that a buyer cannot give a seller more assets than are available, which here means \( \hat{a}_2 \)). Notice \( \lambda(q) > 0 \) because a buyer never brings more than necessary as he can be robbed – intuitively, if he were to have some \( \hat{a}_2 \) left after a purchase, e.g., lowering \( \hat{a}_2 \) would be a profitable deviation as long as \( \delta > 0 \).

In any case, the Euler equations are:

\[ 0 = \hat{a}_{1,t} \left[ \beta (\phi_{t+1} + \rho) - \phi_t \right] \]
\[ 0 = \hat{a}_{2,t} \left\{ \beta (\phi_{t+1} + \rho) (1 - \delta) [1 + \alpha \lambda(q_{t+1})] - \phi_t \right\} \]

Normalize the aggregate asset supply to 1. Then the dynamical system implied by the model is derived as follows. At any date \( t \), there are three possibilities or regimes: (i) \( \hat{a}_{2,t} = 0 \); (ii) \( 0 < \hat{a}_{2,t} < 1 \); and (iii) \( \hat{a}_{2,t} = 1 \). In regime (i), inserting \( \hat{a}_{1,t} = 1 \) and \( \hat{a}_{2,t} = 0 \) into (32) and (33), we get \( \phi_t = \beta(\phi_{t+1} + \rho) \) and \( (1 - \delta) [1 + \alpha \lambda(q_{t+1})] \leq 1 \), with the latter equivalent to

\[ \delta \geq \hat{\delta} \equiv \frac{\alpha \lambda(0)}{1 + \alpha \lambda(0)}. \]

Thus, agents hold no asset in the DM if the probability of theft is high. With Kalai bargaining and the Inada condition \( u'(0)/v'(0) = \infty \), this reduces to \( \hat{\delta} = \alpha \theta/(1 - \theta + \alpha \theta) \), so \( \hat{\delta} = 1 \) if \( \theta = 1 \) and \( \hat{\delta} < 1 \) otherwise. If (34) holds, the DM shuts down, in which case it should be obvious that the only possible outcome is \( \phi_t = \phi^F \)

\( \forall t \), where \( \phi^F \equiv \beta \rho/(1 - \beta) \) is what we call the fundamental price of the asset.\(^{22}\)

\(^{22}\)One might argue that \( \phi^F \) should solve \( \phi = \beta(1 - \delta)(\phi + \rho) \), not \( \phi = \beta(\phi + \rho) \), since the
Now assume the theft probability is low, \( \delta < \tilde{\delta} \), and consider regime (ii), where agents hold some but not all their assets as liquid asset. Inserting \( \hat{a}_{1,t}, \hat{a}_{2,t} > 0 \) into (32) and (33), we get \( \phi_t = \beta (\phi_{t+1} + \rho) \) and \( (1 - \delta) [1 + \alpha \lambda (q_{t+1})] = 1 \), which means \( q_{t+1} = \tilde{q} \) where

\[
\alpha \lambda (\tilde{q}) = \frac{\delta}{1 - \delta}.
\]

Again, since \( \lambda \) is decreasing (see Gu and Wright 2016), \( \tilde{q} \) is the highest possible \( q \) one can get given the probability that asset can be stolen. Thus, \( \hat{a}_{2,t} < 1 \) obtains iff \( \phi_{t+1} + \rho > \hat{a}_{2,t} (\phi_{t+1} + \rho) = v(\tilde{q}) \), as well as \( \delta < \tilde{\delta} \), as assumed so the DM can operate.

Finally, consider regime (iii). Inserting \( \hat{a}_{1,t} = 0 \) and \( \hat{a}_{2,t} = 1 \) into (32) and (33), we now get \( \phi_t \geq \beta (\phi_{t+1} + \rho) \) and

\[
\phi_t = \beta (\phi_{t+1} + \rho) (1 - \delta) [1 + \alpha \lambda (q_{t+1})],
\]

where \( q_{t+1} = v^{-1}(\phi_{t+1} + \rho) < \tilde{q} \). This last condition is equivalent to \( \phi_{t+1} \leq \tilde{\phi} \equiv v(\tilde{q}) - \rho \). Hence if \( \delta < \tilde{\delta} \) the dynamic system is \( \phi_t = \Phi(\phi_{t+1}) \) where:

\[
\Phi(\phi) \equiv \begin{cases} 
\beta (\phi + \rho) (1 - \delta) [1 + \alpha \lambda \circ v^{-1}(\phi + \rho)] & \text{if } \phi < \tilde{\phi} \\
\beta (\phi + \rho) & \text{if } \phi \geq \tilde{\phi}
\end{cases}
\]

This says if the asset price is low, all assets will be used in the DM. Otherwise, agents will hold some assets in illiquid form.

As usual, equilibrium is a nonnegative and bounded path for \( \phi_t = \Phi(\phi_{t+1}) \). A steady state solves \( \tilde{\phi} = \Phi(\tilde{\phi}) \). It is easy to characterize it, in terms of the same three regimes of steady states, as shown in Fig. 10.

---

asset holder only gets the return when the asset is not stolen. A rebuttal is that someone gets the payoff, even if it is only the thief. This issue is purely semantic, and one can simply interpret \( \phi^F \) as notation for \( \beta \rho / (1 - \beta) \).
Proposition 4 A steady state exists, is unique, and is described as follows. Define \( \tilde{\delta} \in [0, \hat{\delta}) \) by

\[
\tilde{\delta} = \frac{\alpha \lambda \circ \nu^{-1}(\phi F + \rho)}{1 + \alpha \lambda \circ \nu^{-1}(\phi F + \rho)}.
\] (38)

Then (i) \( \delta \geq \tilde{\delta} \) implies \( \hat{a}_1 = 1, \hat{a}_2 = 0 \) and \( \tilde{\phi} = \phi F \); (ii) \( \delta \in (\tilde{\delta}, \hat{\delta}) \) implies \( \hat{a}_1 > 0, \hat{a}_2 > 0 \) and \( \tilde{\phi} = \phi F \); and (iii) \( \delta \leq \tilde{\delta} \) implies \( \hat{a}_1 = 0, \hat{a}_2 = 1 \) and \( \tilde{\phi} > \phi F \).

In a regime (i) steady state the DM is inactive and \( \tilde{\phi} = \phi F \), because assets are not safe enough to use as DM payment instruments.\(^{23}\) In regime (ii) the DM is active at \( q = \tilde{q} > 0 \), but since there is still some wealth parked in the illiquid \( \hat{a}_1 \), we again have \( \tilde{\phi} = \phi F \). As Fig. 10 shows, in regime (ii) \( \partial q / \partial \delta < 0 \), because \( \lambda(q) \) is decreasing in \( q \). Thus DM output goes down with \( \delta \) because it reduces output per trade, \( \bar{q} \), as well as the number of trades, \( 1 - \delta \). The regime (iii) steady state is the most interesting, perhaps, since it maximizes ‘cash in the market’ with \( \hat{a}_2 = 1 \), and implies for \( \tilde{\phi} > \phi F \) due to the liquidity premium. In this regime \( \partial q / \partial \delta < 0 \) not because \( \hat{a}_2 \) falls but because \( \phi \) falls with \( \delta \). Also notice that \( \tilde{\delta} \) is decreasing in \( \rho \), so for a given \( \delta \) the economy is more likely to be in regime (iii) when \( \rho \) is smaller, which means liquidity is relatively scarce.

\(^{23}\)The DM can be active in regime (i) if we allow some barter or some credit, as is easy to do, but still assets would not be used in payments, and so again \( \phi_t = \phi^F \forall t \).
As Fig. 11 shows, the steady state can either be on the linear or the nonlinear branch of the system $\Phi(\cdot)$. In the former case, $\bar{\phi} = \phi^F > \tilde{\phi}$ and the steady is the only equilibrium (any other path for $\phi_t$ is unbounded). In the latter case, $\phi^F < \tilde{\phi} < \bar{\phi}$, and we could have cyclic, chaotic or stochastic equilibria where, $\phi_t$ oscillates around $\tilde{\phi}$, which as usual happens when $F'(\tilde{\phi}) < -1$. So the economy may have asset prices above their fundamental value and they exhibit excess volatility, but only if $\hat{a}_2 = 1$ in steady state, which requires $\delta < \tilde{\delta}$ and hence $\rho$ not too big.

In addition to $a_1$ and $a_2$, let us now consider an alternative way to hold the asset. Suppose agents can deposit asset at a bank in exchange for a bank note. Bank note is safer, if not completely safe, than holding $a_2$.\textsuperscript{24} Its return may be lower than $\rho$ as the bank may incur a cost to manage deposit. Let $\iota$ be the return on bank notes. Let $a_3$ denote the amount of deposits. Bank’s profit from having assets as deposits is

$$\Pi(a_3) = a_3 (\rho - \iota) - k(a_3),$$

where $k(a_3), k'(a_3), k''(a_3) \geq k(0) = 0$ is the cost of managing the funds. Profit maximization equates the spread $\rho - \iota$ to the marginal cost $k'(a_3)$.

It is possible to have $\hat{a}_j > 0 \ \forall j$, in general, because assets have heterogeneous attributes. Moreover, buyers may want to hold both $\hat{a}_2$ and $\hat{a}_3$ for diversification – holding one partially insures against the loss of the other. However, for our analysis, it is sufficient to assume that bank notes are perfectly safe and there is no cost of managing funds. Therefore, holding $a_3$ strictly dominates holding $a_2$. Under these assumptions, the economy looks similar to the one without banking except, we replace $\delta$ with 0. In terms of Fig. 11, this shows up as an upward shift in the nonlinear branch of $\Phi(\cdot)$. In general, this shift increases $\bar{\phi}$, as well as $\tilde{\phi}$, naturally, since agents can now park their assets in a safe place – the bank – and still be able to pay in the DM. This is precisely what banking achieves, as discussed in the historical, if not the theoretical, literature. Notice that allowing banks, like reducing $\delta$ in the economy without banks, increases DM output both because it increases output per

\textsuperscript{24}This can be interpreted as the bank may abscond with deposits with some probability.
trade and the number of trades. But what does it do for volatility?

Start without banks, and suppose the steady state is on the linear branch of \( \Phi(\cdot) \), so there is a unique equilibrium and \( \phi_t = \phi^F \ \forall t \). Then introduce banks. This could shift the nonlinear branch of \( \Phi(\cdot) \) up by enough that the new steady state is now on the nonlinear branch. Then it is obvious that the introduction of banking makes possible cyclic, chaotic or stochastic equilibria that were impossible without banking. As mentioned, it could be that these phenomena were possible without banking, but the following is now clear: if the economy has a unique equilibrium with banking where \( \phi = \phi^F \), the same is true without banking. We summarize as follows:

**Proposition 5** Consider an economy without banking that has a unique equilibrium, the steady state with \( \bar{\phi} = \phi^F \). (i) If we introduce banking it is possible to get \( \bar{\phi} > \phi^F \), which is good for DM trade in steady state, but can introduce new nonstationary equilibria. (ii) When the economy with banking has a unique equilibrium, it is a steady state with \( \bar{\phi} = \phi^F \), the same as the economy without banking.

Part (i) of the above result says that banks may engender volatility in a precise sense. Part (ii) says that in a sense they cannot eliminate volatility, since if there is a unique equilibrium with \( \phi_t = \phi^F \ \forall t \) after introducing banking, there is also a unique equilibrium with \( \phi_t = \phi^F \ \forall t \) before banking. As an explicit example, consider

\[
u(q) = \frac{A}{1-\sigma} \left[ (q + \varepsilon)^{1-\sigma} - \varepsilon^{1-\sigma} \right], \quad c(q) = q,
\]

and bargaining with \( \theta = 1 \).

**Example 7:** Let \( A = 0.15, \sigma = 3.1, \varepsilon = 0.16, \rho = 0.033, \beta = 0.8333, \delta = 0.85 \) and \( \alpha = 1 \).

Without banks, this economy has a unique equilibrium and it is the steady state with \( \bar{\phi} = \phi^F = 0.1650 \). If we now introduce banks it has a steady state of \( \bar{\phi} = 0.3183 \) and a two-cycle of \( \phi_L = 0.3193 \) and \( \phi_H = 0.3502 \). Fig. 11 shows the equilibrium of this example.
While examples suffice to make the point that, under some conditions, banking can contribute to volatility, we also want to relax a few of the assumptions to see what else the model can generate. Here we ask if it can generate concurrent circulation of assets and bank liabilities in the DM. To this end we stick to the limit that bank notes will get stolen with probability 0, but, so that \( \hat{a}_3 \) does not strictly dominate \( \hat{a}_2 \), we assume banks have cost function \( k(a_3) \). Assume \( k' (a_3) > 0 \) and \( k'' (a_3) \geq 0 \ \forall a_3 > 0 \), and impose the Inada conditions \( k'(0) = 0 \) and \( k'(1) = \infty \). Then banks maximize \( \Pi (a_3) \) by issuing deposits until \( \rho - \iota = k'(a_3) \), defining a supply curve that is decreasing in \( \rho - \iota \), with \( a_3 \to 0 \) as \( \iota \to \rho \) and \( a_3 \to 1 \) as \( \iota \to 0 \). In equilibrium \( \iota \) depends on \( a_3 \) although a competitive bank takes \( \iota \) as given.

The equilibrium is characterized by equations (32)-(33) with

\[
q_{t+1} = v^{-1} \left\{ (\phi_{t+1} + \rho) \hat{a}_2 + [\phi_{t+1} + \iota (\hat{a}_3)] \hat{a}_3 \right\},
\]

which is the purchase in DM using both \( a_2 \) and \( a_3 \), and

\[
0 = \hat{a}_{3, t} \left\{ \beta (\phi_{t+1} + \iota_t) \left[ 1 + \delta \alpha \lambda (q_{t+1}') + (1 - \delta) \alpha \lambda (q_{t+1}) \right] - \phi_t \right\} \tag{40}
\]

where \( q_{t+1}' = v^{-1} \left\{ (\phi_{t+1} + \iota (\hat{a}_3)) \hat{a}_3 \right\} \) denotes the purchase when \( a_2 \) gets stolen. It is easy to show there exists a steady state in which both assets and bank notes are
used as means of payments. Moreover, the following example gets cycles in which
the issuance of bank notes fluctuates endogenously with changes in price.

**Example 8:** Let \( A = 2.5, \sigma = 2.5, \varepsilon = 0.001, \rho = 0.04, \beta = 0.8, \delta = 0.01 \) and \( \alpha = 1. \) For bank’s technology, assume \( k' = 0.03 \forall a_3 \) so \( \upsilon = 0.01. \)

There is a unique steady state in which \( \bar{\phi} = 1.3125 \) and \( \bar{a} = (0, 0, 1). \) There
also exists a two-cycle in which \( \phi_L = 1.2128, \hat{a}_L = (0, 0, 1), \phi_H = 1.4760 \) and
\( \hat{a}_H = (0.0384, 0.2293, 0.7323). \) The price of asset fluctuates around the steady state.
However, the value of deposit \( (\phi_t + \iota_t) a_{3,t} \) is smaller in both periods of a cycle.

Fig. 12 plots the time series of asset price \( \phi, \) the amount of deposit \( a_3, \) value of de-
posit \( (\phi + \iota) a_3, \) and trade surplus at \( t \) in this example. When \( \phi \) is low, the return on
asset is high, and agents can afford to purchase more deposit. The value of deposit
moves with \( a_3 \) in this example even though \( \phi \) decreases with \( a_3. \) The trade surplus
moves against \( a_3 \) as when \( a_2 \) is not stolen, the total liquidity is higher, so is the expected surplus.

![Figure 12: Model 4A, time series for a two-period cycle](image)

At this point we change the nature of safety to accommodate the idea of banks as
secret keepers, as in literature mentioned above, although the way we do it is slightly
different. Namely, as in several related papers (e.g., Rocheteau and Rodriguez-Lopez
2014; Hu and Rocheteau 2015; Lagos and Zhang 2019), we assume assets survive
with probability $\delta$ but disappear otherwise at the beginning of each CM, where they are replaced by new units, to maintain stationarity, distributed lump-sum across agents. We assume this shock is aggregate (all or no assets depreciate), so one cannot insure it away by holding a diversified portfolio, and that information about the shock in the next CM is revealed in the current DM. The shocks are therefore a hindrance to assets serving as a media of exchange, and while it is extreme, in principle all we really need is aggregate socks to asset values, but they need not drop to 0.

In any case, the CM problem is

$$W_t(a_t) = \max_{x_t, t, \hat{a}_t} \{ U(x_t) - \ell_t + \beta V_{t+1}(\hat{a}_t) \} \text{ st } x_t = (\phi_t + \rho) a_t + \ell_t - \phi_t \hat{a}_t + T$$

where $T$ is the transfer. Now, as the asset is the only DM means of payment, and it is only used if it survives in the next CM, we have

$$V_{t+1}(\hat{a}_t) = \delta \alpha [u(q_{t+1}) - v(q_{t+1})] + \delta W_{t+1}(\hat{a}_t) + (1 - \delta) W_{t+1}(0)$$

where $v(q_{t+1}) = (\phi_{t+1} + \rho) \hat{a}_t$ if $(\phi_{t+1} + \rho) \hat{a}_t < v(q^*)$ and $v(q_{t+1}) = v(q^*)$ otherwise.

Normalizing the asset supply to 1, as above we derive dynamical system $\phi_t = \Phi_0(\phi_{t+1})$, where

$$\Phi_0(\phi) = \beta \delta (\phi + \rho) \left[ 1 + \alpha \lambda \circ v^{-1}(\phi + \rho) \right]. \quad (41)$$

Let $z = \phi + \rho$ when the asset is not destroyed. Then from (41) we can get $z_t = f_0(z_{t+1})$, where

$$f_0(z) = \beta \delta z \left[ 1 + \alpha \lambda \circ v^{-1}(z) \right] + \rho$$

Suppose the banks have a costless technology to prevent the information from being revealed in the DM. They issue bank notes in the CM in exchange for asset. Bank notes are redeemed in the next CM. Using bank notes enables agents to trade in the DM using the expected value of the asset instead of the realized value. Since $W_t(\cdot)$ is linear, the DM value thus becomes

$$V_{t+1}(\hat{a}_t) = \alpha [u(q_{t+1}) - v(q_{t+1})] + W_{t+1}(\delta \hat{a}_t)$$
where \( v(q_{t+1}) = \delta (\phi_{t+1} + \rho) \hat{a}_t \) if \( \delta (\phi_{t+1} + \rho) \hat{a}_t < v(q^*) \) and \( v(q_{t+1}) = v(q^*) \) otherwise. Notice that \( V \) looks as if the bank can perfectly diversify asset portfolio to ensure \( \delta \) fraction of the investment will survive. This is true for the DM transaction as agents use the expected value of the asset to trade. In the CM however, agents still encounter an aggregate shock but the linearity of \( W \) implies that their expected utility is the utility of the expected value. With the asset market clearing condition, the equilibrium path of \( \phi_t \) derived from the Euler equation is 

\[
\Phi_1 (\phi) = \beta \delta (\phi + \rho) \left\{ 1 + \alpha \lambda \circ v^{-1} [\delta (\phi + \rho)] \right\} .
\]

(42)

Let \( z = \delta (\phi + \rho) \) be the liquidity used in the DM. Then from (42) we get 

\[
z_t = f_1 (z_{t+1}),
\]

where

\[
f_1 (z) = \beta \delta z \left[ 1 + \alpha \lambda \circ v^{-1} (z) \right] + \delta \rho
\]

Again \( \lambda (q) \) is decreasing, so \( \Phi_1 \) lies above \( \Phi_0 \), meaning the asset price is higher in steady state with bank notes. It is obvious that \( f_0 \) is an upward shift of \( f_1 \), meaning that the liquidity provided by bank notes in steady state is lower than that provided by the asset if not destroyed. Welfare can be improved using bank notes, as they provide a steady stream of liquidity, but they also can engender instability.

**Example 9** Consider the functions and the terms of trade in Examples 7 and 8. Let \( A = 0.75, \sigma = 3.5, \varepsilon = 0.15, \rho = 0.85, \beta = 0.9, \delta = 0.1 \) and \( \alpha = 1 \).

Fig.13 shows the equilibrium with and without banking. In the steady state without banking, \( z = 0.9341 \) and \( q = q^* = 0.7711 \) if asset survives, and \( z = q = 0 \) if it does not. Indeed, the steady state is the only equilibrium in this economy. With banking, \( z = q = 0.3511 \) in all periods. The expected surplus without banking is 3.3287, while with banking it is 32.3876. However, with banking, there exists a two-period cycle in which \( z_L = 0.1625 \) and \( z_H = 0.7280 \). The time series of two-cycle is simple here so we do not display it: all variables including \( \phi, q, \) and trading surplus move with \( z \). Thus, banking eliminates fundamental cycles induced by information on asset, but it introduces self fulfilling cycles induced by beliefs. Hence, we state the following proposition.
Proposition 6  Consider an economy without banking that has a unique equilibrium, the steady state with $z > v(q^*)$. If we introduce banking it is possible to get $z < v(q^*)$ and can introduce new nonstationary equilibria.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{model4b.png}
\caption{Model 4B, dynamic equilibrium}
\end{figure}

Conclusion

As discussed in our introductory remarks, there is a well-documented view that financial intermediaries are in some way inherently unstable or fragile. Our approach to exploring this hypothesis involved building formal models of these institutions and asking if they make it more likely that there can be multiple equilibria and excess volatility. It seems they do. We also wanted to know if the answer depends on how one chooses to focus on particular aspects of their many and varied activities. Therefore, we considered four distinct settings. Two involved reputation or trust – something that seems quite at home in a theory of banking – but those models differed in the fundamental reason for the emergence of banking: one concerned insurance; the other emphasized fixed costs of investment projects. Another setup was not designed to capture banks, but other types of financial intermediaries, namely middlemen in OTC asset markets. Yet another concerned the use of bank liabil-
ities in the payment system, which is commonly thought to be one of the most distinguishing features of banking, even though it shows up less in the mainstream literature on banking than in, say, monetary theory.

Although the models are different in several economic and mathematical respects, they have some commonalities, and they display similar results related to multiplicity and volatility.\textsuperscript{25} The fact that these substantive results more or less transcend details of formalization seems to us to lend credence to the notion that financial intermediation may indeed engender instability. But it cannot be emphasized strongly enough that this does not imply financial intermediation is a bad idea.\textsuperscript{26} Also, of course, many of our results came in the form of examples, where we made little effort to calibrate parameters realistically, and in any case, to reiterate something said earlier, the claim is not that actual data are best explained by cycles or sunspots. The suggestion instead is this: if our rudimentary models of financial intermediation can have equilibria where liquidity, prices, quantities and other variables vary as self-fulfilling prophecies, maybe intermediated asset markets in the real world might be prone to similar instability.

During the project we developed some new theoretical environments and combined others in novel ways (e.g., extending a typical model in the Diamond-Dybvig tradition to be a little more dynamic), and perhaps others will find this useful in their applications. Moreover, we tried to have financial intermediaries arise endogenously, in other words to study models of the institutions and not merely models with them. This meant working with economies with explicit frictions, like limited commitment, spatial or temporal separation, imperfect information or communication, etc. Heuristically, this helps explain some of the results: In these settings, in contrast to classsical GE theory à la Debreu, standard concepts of equilibrium do

\textsuperscript{25}Rajan (2005) attributes the fluctuations in the financial sector to moral hazard problem. Whereas our paper emphasizes that the exact friction that necessitates financial intermediation endangers stability of the market.

\textsuperscript{26}Summers in his discussion of Rajan (2005) argues that the innovation in financial sector, like the development in the transportation technology, has a overwelmingly positive effect on the world, although just as there are traffic accidents, there is some negative effect such as financial crisis as well.
not necessarily deliver efficiency, which is what provides room for institutions like money, banks, etc. This is what ties together our various formalizations. Moreover, it ties in with a conjecture Shell (1992) calls The Philadelphia Pholk Theorem, which says that in all models where equilibria are not efficient one can find results like the ones we highlight, like multiple or sunspot equilibria. It is hard to prove this, because it concerns all models, not particular models, but our findings are consistent with the idea.
Appendix A: Proofs of Nonobvious Results

Proposition 1: If $\hat{\phi} \leq 0$ then (12) reduces to $\phi_{t-1} = \beta \phi_t$ and the only equilibrium is the steady state with $\bar{\phi} = 0$. If $\hat{\phi} > 0$ then $f(0) > 0$ and $f(\hat{\phi}) = \beta \hat{\phi} < \hat{\phi}$ implies $\bar{\phi} \in (0, \hat{\phi})$ exists. To see it is unique, first solve (12) for $\phi = \beta \mu v(bR) / (1 - \beta)$ and substitute it into (10) to get $\lambda d = (1 - \beta + \beta \mu) v(bR) / (1 - \beta)$. This implies $d$ is increasing in $b$. But (11) implies $d$ is decreasing in $b$, so if they have a solution $(\bar{b}, d)$ it is unique. ■

Proposition 3: First, for uniqueness, note that when $\pi = \bar{\pi}$ the equations for the $R$-curve and $n$-curve are defined by

$$
\frac{(r + \delta + \alpha \theta_{ms})}{1 - \delta} R - \rho = \frac{\alpha m_b \theta_{mb} (\bar{\pi} - R) + \alpha (n_b + n_m) \theta_{ms} R}{N} = 0 \quad (43)
$$

$$
\delta n + n (1 - \delta) \frac{\alpha m_b \tau}{N} - (n_m - n) (1 - \delta) \frac{\alpha \left(1 - \frac{n_b + n_m}{N}\right)}{(1 - \delta)} = 0 \quad (44)
$$

where

$$
N = \frac{\alpha m_b \theta_{sa} n + \alpha (n_m - n) \theta_{sn} R}{\kappa_s}.
$$

In the region where $R > \bar{\pi}$, where $\tau = 0$, combine (43) and (44) to eliminate $N$,

$$
\left(\frac{r + \delta + \theta_{ms} \delta n}{n_m - n}\right) R = \rho (1 - \delta)
$$

(45)

This implies

$$
\frac{\partial R}{\partial n} = -\frac{\theta_{ms} \delta n_m}{(n_m - n)^2(r + \delta) + (n_m - n) \theta_{ms} \delta n} < 0.
$$

Thus we transform the system (43)-(44) to (45)-(44). As (45) is downward sloping and (44) is upward sloping, there exists at most one steady state with $R > \bar{\pi}$.

In the region where $R < \bar{\pi}$, where $\tau = 1$, combine (43) and (44) to get

$$
\left(\frac{r + \delta + \alpha \theta_{ms}}{1 - \delta}\right) R = \rho + \frac{n_b \theta_{mb} (\bar{\pi} - R) + (n_b + n_m) \theta_{ms} R}{(1 - \delta) n_m (n_b + n_m - n)} [(n_m - n) (1 - \delta) \alpha - n\delta].
$$

This implies

$$
\frac{\partial R}{\partial n} = -\frac{n_b \theta_{mb} (\bar{\pi} - R) + (n_b + n_m) \theta_{ms} R \delta (n_b + n_m) + (1 - \delta) \alpha n_b}{r + \delta + (\theta_{ms} + n_b \theta_{mb})(1 - \delta) \alpha/N n_m (n_b + n_m - n)^2} < 0 \quad (46)
$$

Again, since (46) is downward and (44) upward sloping, there is at most one steady state with $R < \pi$. Similarly, when $R = \bar{\pi}$ and $\tau \in (0, 1)$, the $n$-curve is flat and $R$-curve is upward sloping. Hence, there again is at most one steady state.

39
For existence, first, it is easily verified that the $R$- and $n$-curve are upward sloping. At $n = 0$ the $R$-curve implies $R > 0$ and the $n$-curve implies $n > 0$. At $R = \infty$ the $R$-curve implies $n = n_m$ and the $n$-curve implies $n = \bar{n} \equiv \alpha n_m (1 - \delta) / [\delta + (1 - \delta) \alpha] < n_m$. Hence the curves cross at least once, and generically an odd number of times. Since we already established that there cannot be multiple steady states in the same regime, if there is a steady state at $R = \bar{R}$, there must exist two other steady states, one with $R < \bar{R}$ and one with $R > \bar{R}$. Routine calculation implies $\partial R/\partial \rho > 0$, and so there exist $\bar{\rho}, \bar{\rho} \geq 0$ with the properties specified in Proposition 3. ■

Proposition 7  First note that $G(0) = \infty$, so $G(R)$ starts above the 45° line. Also, $G(R) = \beta (1 - \delta) [\rho + R - (1 - \beta) \kappa_m]$ for $R$ above the upper bound of $\pi$, so it is linear with $G'(R) < 1$ when $R$ is big. Hence steady state exists. From (54) we derive

$$
\Phi'(R) = \beta (1 - \delta) \left\{ 1 - \frac{(1 - \beta) \kappa_m n_m \theta_m R}{n_s \theta_m s R} \left[ \int_R^{1} \frac{1 - F(\pi)}{R} d\pi + 1 - F(R) \right] \right\} < 1
$$

Therefore, steady state is unique. If $\Phi'(R) < -1$, there exists two-period cycle and thus sunspot equilibria. ■

Appendix B: Sunspot Equilibria

If a dynamical system allows for a two-state stationary sunspot equilibrium, is solves

$$
\phi_{s,t-1} = \rho_s f(\phi_{s,t}) + (1 - \rho_s) f(\phi_{s,t})
$$

where $s = A, B$ denotes two states in the sunspot equilibrium, $\rho_s \in (0,1)$ is the probability of staying in the same state, and $f$ is the function of the dynamical system in the deterministic case (i.e., (12) in Model 1 and (19) in Model 2). We seek a pair of probabilities $(\rho_A, \rho_B) \in (0,1)^2$ satisfying (47). Rewrite (47) as

$$
\rho_A = \frac{f(\phi_B) - \phi_A}{f(\phi_B) - f(\phi_A)} \quad \text{and} \quad \rho_B = \frac{\phi_B - f(\phi_A)}{f(\phi_B) - f(\phi_A)}
$$

Consider wlog $\phi_B > \phi_A$. If $f$ is decreasing on $(\phi_A, \phi_B)$, the denominator is negative. The necessary and sufficient condition for $\rho_A, \rho_B \in (0,1)$ is $f(\phi_A) > \phi_B > \phi_A > f(\phi_B)$, which implies that $f$ crosses the 45 degree line from above and $[f(\phi_A) - f(\phi_B)] / (\phi_A - \phi_B) < -1$. Therefore, in Model 1 where $f$ is decreasing around the steady state, there exist sunspot equilibria if $f(\phi) < -1$. 40
Similarly, if $f$ is increasing on $(\phi_A, \phi_B)$, the denominator is positive. The necessary and sufficient condition for $\rho_A, \rho_B \in (0, 1)$ is $f(\phi_B) > \phi_B > \phi_A > f(\phi_A)$, which implies $f$ crosses the 45 degree line from below on $[\phi_A, \phi_B]$. Therefore, in Model 2 where $f$ is increasing, there exist sunspot equilibria around the stable steady state $\phi_1$, and any $\phi_A \in (0, \phi_1)$ and $\phi_B \in (\phi_1, \phi_2)$ satisfy the condition and constitute a two-state stationary sunspot equilibria.

Appendix C: Bargaining in Model 2

There are two agents on each island, one who lives for one period and one who lives forever, so the former should be the depositor and the latter the banker. Assume the cost $\kappa$ is too high for them to invest individually. If the banker’s bargaining power is $\theta$, the generalized Nash problem is

$$W(\phi) = \max_{X,x,D,d} \left[ U(X) - C(D) \right]^{\theta} \left[ u(X) - c(d) \right]^{1-\theta}$$

$$\text{st } X + x = R(D + d) - \kappa$$

$$u(x) - c(d) \geq 0$$

$$x_t \leq \phi_t.$$  \hspace{1cm} (48)

As usual the last constraint from the banker’s incentive condition $\beta V_{t+1} \geq \lambda x_t + (1 - \mu) \beta V_{t+1}$ rewritten using $\phi_t \equiv \beta u V_t / \lambda$. Notice $W'(\phi) > 0$ if (51) binds, and there is a cutoff $\phi$ above which banking is viable and below which it is not.

Denote the solution ignoring (50) and (51) by $(X^*, x^*, D^*, d^*)$. Further, consider the case $u(x^*) > c(d^*)$ and let $\phi^* = x^*$. Substituting (49) into the objective function and taking FOC’s, we get

$$D : U'(X) R - C'(D) = 0$$

$$d : \theta U'(X) R[u(x) - c(d)] - (1 - \theta) c'(d) [U(X) - C(D)] - \eta_1 c'(d) = 0$$

$$x : -\theta U'(X) [u(x) - c(d)] + (1 - \theta) u'(x) [U(X) - C(D)] + \eta_1 u'(x) - \eta_2 = 0$$

where $\eta_1$ and $\eta_2$ are multipliers. From this one can see the banker’s surplus may not increase with $\phi$ at least close to $\phi^*$:

$$\frac{\partial^2 [U(X) - C(D)]}{\partial \phi} \bigg|_{\phi = \phi^*} = \frac{R^2 U''(d)}{C''(c^*)} (1 - \theta) c''(d) < 0$$

The banker’s value function is $V_t = U(X_t) - C(D_t) + \beta V_{t+1}$, and using $\phi_t = \beta u V_{t+1} / \lambda$ we have

$$\phi_{t-1} = \frac{\beta \mu}{\lambda} [U(X_t) - C(D_t)] + \beta \phi_t.$$  \hspace{1cm} (52)
Now (52) can be written as

\[
\phi_{t-1} = \begin{cases} 
\beta \phi_t & \text{if } \phi_t < \bar{\phi} \\
\frac{\beta \mu}{\lambda} [U \circ X (\phi_t) - C \circ D (\phi_t)] + \beta \phi_t & \text{if } \bar{\phi} \leq \phi_t < \phi^* \\
\frac{\beta \mu}{\lambda} [U (X^*) - C(D^*)] + \beta \phi_t & \text{if } \phi_t \geq \phi^*
\end{cases}
\]

Fig. AC: Delegated monitoring – Nash bargaining

Fig. AC shows the dynamical system for the following parameterization:

**Example AC:** Let \( U(x) = u(x) = Ax \) and \( C(d) = c(d) = Bd^\gamma/\gamma \), where \( A = 1, B = 0.5, \gamma = 5, R = 2, k = 1.5, \theta = 0.01, \lambda = 0.01, \mu = 1 \) and \( \beta = 0.35 \).

There are three steady states, \( \phi = 0 \) and \( \phi_2 > \phi_1 > 0 \), with \( f \) crossing the 45° line first from below at \( \phi_1 \) and from above at \( \phi_2 \). Hence there are sunspot equilibria around \( \phi_1 \) fluctuating between any \( \phi_A \in (0, \phi_1) \) and \( \phi_B \in (\phi_1, \phi_2) \), similar to the baseline version of Model 2, and since \( f'(\phi_2) < -1 \) so there is a two-cycle with periodic points \( \phi_L \) and \( \phi_H \), plus sunspot equilibria for any \( \phi_A \in (\phi_L, \phi_2) \) and \( \phi_B \in (\phi_2, \phi_H) \), similar to Model 1.

**Appendix D: Entry by M in Model 3**

Here we consider entry by \( M \) with \( n_b \) and \( n_s \) fixed. The equations (23)-(27) are the same with entry by \( M \), except now \( n_s \) is fixed while \( n_{m,t} \) is endogenous. What changes is the participation condition: \( V_{s,t} = \kappa_s \) is replaced by \( V_{0,t} = \kappa_m \), assuming \( M \) acquires inventory after entry. Given this, (24) yields a simple expression for \( N_t \).
in terms of $R_t$,

$$N_t = \frac{\alpha n_s \theta_{ms} R_t}{(1 - \beta) \kappa_m}. \quad (53)$$

From (53), $N_t$ depends only on $R_t$, while in other version, with entry by $S$, it depends on $R_t$ and $n_t$. Substituting (53) into (26), after some algebra we get $R_{t-1} = G(R_t)$, where

$$G(R) \equiv \beta (1 - \delta) \left\{ \rho + R + \frac{(1 - \beta) \kappa_m n_b \theta_{mb}}{n_s \theta_{ms} R} \int_R^\infty [1 - F(\pi)] d\pi - (1 - \beta) \kappa_m \right\}. \quad (54)$$

From (54), $R_{t-1}$ depends only on $R_t$, while in other version, it depends on $R_t$ and $n_t$.

These considerations make entry by $M$ easier, since it delivers the univariate system $R_{t-1} = G(R_t)$, which determine the path for $R_t$, after which $N_t$ follows from (53), $n_t$ from (27), etc. If fact, given a fixed point $R = G(R)$, we have

$$N = \frac{\alpha n_s \theta_{ms} R_t}{(1 - \beta) \kappa_m},$$

$$n_m = \frac{\alpha n_s \theta_{ms} R_t}{(1 - \beta) \kappa_m} - n_s - n_b,$$

$$n = \frac{n_s (1 - \delta) [\alpha n_s \theta_{ms} R - (n_b + n_s) (1 - \beta) \kappa_m]}{\delta n_s \theta_{ms} R + (1 - \delta) [n_b \tau (R) + n_s] (1 - \beta) \kappa_m}.$$

To guarantee the fixed point is a steady state we must check $n_m, n \geq 0$, both of hold iff $R \geq R \equiv (n_s + n_b) (1 - \beta) \kappa_m / \alpha n_s \theta_{ms}$ (we also need $n \leq n_m$ but that never binds). Hence, a solution to $R = G(R) \geq R$ is a steady state with type $M$ active; if there is we no such $R$, the only steady state has no intermediation.

![Figure AD: Model 3, cycles with entry by $M$](image)
One can check $G(0) = \infty$, $G'(R) < 1$ and $G''(R) \geq 0$. Also, $G$ is linear with slope $\beta(1 - \delta) \forall R > \max(\pi)$. This is shown in the left panel of Fig. AD, from which it is clear that there exists a unique fixed point, say $\hat{R}$. For what it’s worth, we can borrow with minor modification the intuition from Rocheteau and Wright (2005) for uniqueness in this version, vis a vis multiplicity in the baseline Model 3 with entry by $S$: with entry by $S$ there are strategic complementarities, because $M$ choosing to trade more with $B$ encourages entry by $S$, and vice versa; that is absent with entry by $M$, because the agents choosing whether to trade are the same ones choosing to enter. In any case, we can have $\hat{R} > \max(\pi)$ on the linear increasing part of $G(R)$ or $\hat{R} < \max(\pi)$, on the nonlinear part of $G(R)$. If $G'(-\hat{R}) < -1$ then $\hat{R}$ is locally stable, and there exist cycles and sunspots, as in Model 1. Is $G'(\hat{R}) < -1$ possible? Yes, in fact, there is a threshold, say $\rho_1$, such that $G'(\hat{R}) < -1$ if $\rho < \rho_1$. We do not know if $\rho_1 > 0$ or $\rho_1 < 0$, in general, but all our examples gave $\rho_1 < 0$. Still we have to verify $R \geq \hat{R}$ to ensure that $M$ enter the market, as discussed above. It $G'(\hat{R}) < -1$ and $\hat{R} \leq R$ possible? Yes, as we now show by way of example. 

**Example AD:** Consider $\alpha = 1$, $\delta = 0.01$, $\beta = 0.99$, $n_b = n_s = 1$, $\theta_{mb} = \theta_{ms} = 0.5$, $\kappa_m = 0.1$, $\rho = -0.1$, and use the $F(\pi)$ in (30), with $\pi_0 = 4.95$, $\pi_1 = 0.05$, $\pi_2 = 5.05$, $\pi_3 = 0.95$ and $\pi_4 = 10$. This is the example used in Fig. AD, where it can be easily checked that $G'(\hat{R}) < -1$ and $\hat{R} < R$. Hence this example admits a two-cycle. Note that $\rho < 0$ in this example. If we lower $\rho$ a little more, we can get higher-order cycles, including three-cycles, and hence chaotic dynamics. This is shown in the right panel of Fig. AD, where we plot $G^3(R)$ and see that there exist fixed points other than $\hat{R}$, namely a pair of three cycles. In our examples the dynamics do not involve regime switching; over the cycle, $M$ and $B$ always trade. The dynamics over a two cycle are similar to Fig. 9, from the model with entry by type $S$, so we do not repeat that discussion. However, here we can explicitly construct higher order cycles, and an example is shown in Fig. AD. Finally, one more result is that $\rho < 0$ implies $M$ and $B$ must trade for some $\pi$, $Pr(R < \pi) > 0$, if $M$ is in the market – just like in the other version, a buy-and-hold-forever strategy is never a good idea at $\rho < 0$.

We state this formally as follows:

**Proposition 7** Consider Model 3 with entry by $M$. There is a unique steady state. If $\rho < \bar{\rho}$ there exist cyclic and sunspot equilibria. Also, if $\rho < 0$ then $Pr(R < \pi) > 0$. 

44
References


