Financial Fragility

Franklin Allen  
Department of Finance  
Wharton School  
University of Pennsylvania  
Philadelphia, PA 19104

Douglas Gale  
Department of Economics  
New York University  
269 Mercer Street  
New York, NY 10003  
douglas.gale@nyu.edu

January 31, 2002

\footnote{Earlier versions of this paper were presented at the Cowles Foundation Conference on “The Future of American Banking: Historical, Theoretical, and Empirical Perspectives”, and at the University of Iowa. We are grateful to the participants for their comments, particularly Russell Cooper, Douglas Diamond and Stephen Morris. The financial support of the National Science Foundation, the Financial Institutions Center at the University of Pennsylvania, and the C.V. Starr Center for Applied Economics at New York University are gratefully acknowledged.}
Abstract

We study the relationship between sunspot equilibria and fundamental equilibria in a model of financial crises. There are many sunspot equilibria. Only some of these are the limit of fundamental equilibria when the exogenous fundamental uncertainty becomes vanishingly small. We show that under certain conditions the only “robust” equilibria are those in which extrinsic uncertainty gives rise to financial crises with positive probability.

JEL Classification: D5, D8, G2
1 Excess sensitivity and sunspots

There are two traditional views of banking panics. One is that they are spontaneous events, unrelated to changes in the real economy. Historically, panics were attributed to “mob psychology” or “mass hysteria” (see, e.g., Kindleberger (1978)). The modern version of this view, developed by Bryant (1980), Diamond and Dybvig (1983), and others, replaces mob psychology and mass hysteria with an equilibrium model in which crises result from self-fulfilling expectations.

An alternative to this view is that banking panics are a natural outgrowth of the business cycle. An economic downturn will reduce the value of bank assets, raising the possibility that banks are unable to meet their commitments. When depositors receive information about an impending downturn in the cycle, they anticipate financial difficulties in the banking sector and try to withdraw their funds. This attempt precipitates a crisis.

The formal difference between these two views is whether a crisis is generated by intrinsic or extrinsic uncertainty. Intrinsic uncertainty is caused by stochastic fluctuations in the primitives or fundamentals of the economy. Examples would be exogenous shocks that effect liquidity preferences or asset returns. Extrinsic uncertainty, often referred to as sunspots, by definition has no effect on the fundamentals of the economy.¹,²

Whichever view one takes of the causes of financial crises, there is a consensus that financial systems are fragile. The danger of a financial crisis lies in the possibility that it will propagate through the economic system, causing damage disproportionate to the original shock. This notion of financial fragility is most easily seen in the sunspot model. By definition, extrinsic uncertainty has no effect on the fundamentals of the economy. Hence, any impact of extrinsic uncertainty on the equilibrium of the economy is dispro-

¹Strictly speaking, much of the banking literature exploits multiple equilibria without addressing the issue of sunspots. We adopt the sunspots framework here because it encompasses the standard notion of equilibrium and allows us to address the issue of equilibrium selection.

²The theoretical analysis of sunspot equilibria began with the seminal work of Cass and Shell (1982) and Azariadis (1981), which gave rise to two streams of literature. The Cass-Shell paper is most closely related to work in a Walrasian, general-equilibrium framework; the Azariadis paper is most closely related to the macroeconomic dynamics literature. For a useful survey of applications in macroeconomics, see Farmer (1999); for an example of the current literature in the general equilibrium framework see Gottardi and Kajii (1995, 1999).
portionate to the original shock. Financial fragility can also be captured in a model with intrinsic uncertainty. Here financial fragility is interpreted as a situation in which very small shocks can tip the economy over the edge into a full blown crisis. In other words, financial fragility is an extreme case of excess sensitivity.

Empirical studies of 19th- and early 20th-century American experience, such as Gorton (1988), Calomiris and Gorton (1991), and Calomiris and Mason (2000), consider a broad range of evidence and conclude that the data does not support the view that banking panics are spontaneous events. In a series of papers, Allen and Gale (1998, 2000a-d) investigate equilibrium models of bank runs in which financial crises are triggered by exogenous shocks to liquidity demand and asset returns. In this paper, we use the general equilibrium model introduced in Allen and Gale (2000d), henceforth AG, to investigate the relationship between crises generated by intrinsic uncertainty and crises generated by extrinsic uncertainty (sunspots).

Following Diamond and Dybvig (1983), we model liquidity preference by assuming that consumers have stochastic time preferences. There are three dates, $t=0,1,2$, and all consumption occurs at dates 1 and 2. There are two types of consumers, *early consumers*, who only value consumption at date 1 and *late consumers*, who only value consumption at date 2. Consumers are identical at date 0 and learn their true type, “early” or “late”, at the beginning of date 1.

There are two types of investments in this economy, short-term investments, which yield a return after one period and long-term investments, which take two periods to mature. There is a trade-off between liquidity and returns. Long-term investments have a higher yield but take longer to mature.

Banks are modeled as risk-sharing institutions that pool consumers’ endowments and invest them in a portfolio of long- and short-term investments. In exchange for their endowments, banks give consumers a deposit contract that allows a consumer the option of withdrawing either a fixed amount of

\[\text{A number of authors have developed models of banking panics caused by aggregate risk. Wallace (1988; 1990), Chari (1989) and Champ, Smith, and Williamson (1996) extend Diamond and Dybvig (1983) to allow for a random (aggregate) demand for liquidity. Chari and Jagannathan (1988), Jacklin and Bhattacharya (1988), Hellwig (1994), and Alonso (1996) introduce aggregate uncertainty about asset returns, which can be interpreted as business cycle risk. Postlewaite and Vives (1987) have shown how runs can be generated in a model with a unique equilibrium.}\]
consumption at date 1 or a fixed amount of consumption at date 2. This provides individual consumers with insurance against liquidity shocks by intertemporally smoothing the returns paid to depositors.

There are two sources of liquidity for banks. First, they can hold short-term investments, which are by definition liquid. Secondly, banks have access to an asset market at date 1, in which they can liquidate some of their long-term investments. The existence of a market, rather than a technology for physically liquidating long-term investments, is one of the major differences between AG, on the one hand, and Diamond and Dybvig (1983) and most of the subsequent literature, on the other. It plays a crucial role in the analysis.4

As a benchmark, we first consider an economy in which there is no aggregate uncertainty. More precisely, we assume that (a) the total number of early and late consumers is non-stochastic and (b) asset returns are non-stochastic. In other words, there are no exogenous shocks to (aggregate) liquidity demand or asset returns. As Diamond and Dybvig (1983) argued, equilibrium bank runs can occur in the absence of aggregate uncertainty. Since the bank cannot distinguish early from late consumers, it cannot prevent the late consumers from withdrawing their funds at date 1. Whether it is optimal for a late consumer to withdraw at date 1 or date 2 depends on what he expects the other depositors to do: if every other depositor is expected to withdraw at date 1 there will be no assets left to pay the depositor who withdraws at date 2. If only early consumers withdraw at date 1 the late consumers will be better off waiting until date 2. Thus, there are two “equilibria” at date 1, one in which there is a run involving all depositors and one in which only the early consumers withdraw their deposits at date 1.

The Diamond and Dybvig (1983) analysis of panics is essentially a story about multiple equilibria. The attraction of the theory is that it provides a striking example of financial fragility: bank runs occur in the absence of any exogenous shock to liquidity preference or asset returns. The weakness is that it has little predictive power: bank runs are possible, but not inevitable.

Let us say that a run is inessential if the bank has sufficient resources to meet its commitment to early consumers at date 1 and make an equal or greater payment to late consumers at date 2, that is, it can satisfy the incen-

---

4This modeling device has a number of important implications. Ex post, there is no deadweight loss from liquidating assets. However, the incompleteness markets implies that provision of liquidity is ex ante inefficient. By endogenizing the costs of liquidation, we are able to provide a welfare analysis of prudential regulation (see AG for details).
tive and budget constraints. The panics that result from self-fulfilling expectations are all inessential in this sense. AG, by contrast, ignores self-fulfilling prophecies and focuses on essential runs, that is, runs that occur because the banks cannot simultaneously satisfy the budget and incentive constraints. In the sequel, we follow AG and assume that there are no inessential runs: in equilibrium, late consumers are assumed to wait until date 2 to withdraw from the bank as long as it is incentive-compatible for them to do so.

By restricting attention to essential runs, we can reduce the number of equilibria somewhat and thereby increase the predictive power of the theory. However, even in the absence of intrinsic uncertainty, ruling out panics based on self-fulfilling expectations does not rule out multiple equilibria; nor does it rule out equilibria in which crises occur with positive probability; nor does it rule out a role for extrinsic aggregate uncertainty. By crisis we mean a profound drop in the value of asset prices which affects the solvency of a large number of banks and their ability meet their commitments to their depositors. In extreme cases, the banks are forced into liquidation as the result of an essential run, but there may also be crises in which banks avoid default, although their balance sheet is under extreme pressure.

To see how crises can occur in the absence of exogenous shocks to liquidity demand or asset returns, suppose that banks have already made their investments and commitments at date 0 and consider what happens at date 1. The capital market at date 1 allows banks to trade future (date-2) consumption for present (date-1) consumption. Let $p$ denote the price of future consumption in terms of present consumption; if $\rho$ is the interest rate then $p = 1/(1+\rho)$. If the price $p$ is high (the interest rate is low) the present value of the bank’s assets is high. The bank can meet its commitments to the early consumers at date 1 and to the late consumers at date 2 without selling many assets. If the price $p$ is low (the interest rate is high) the value of the bank’s assets is low and in order to meets its commitments to consumers at date 1 it may have to reduce the payment to consumers at date 2. In the extreme case, it may be impossible for the bank to meet its commitments to the early consumers and still have enough left over to make the late consumers willing to wait until date 2 in which case there will be an essential run. In any case a low value of $p$ forces banks to liquidate a larger fraction of their assets. The endogeneity of $p$ gives rise to the possibility of endogenous crises. If $p$ is low, banks will be forced to liquidate a large fraction of their long-term investments in the capital market. The supply of future consumption (in the form of long-term assets) is high relative to the supply of present consump-
tion, so the market will only clear if the price of future consumption is low. Conversely, if \( p \) is high, there is no need for large asset sales, banks will not liquidate their long-term investments, and the capital market will clear at a high value of \( p \).

This suggests the existence of two types of equilibria. In the first type, \( p \) is non-stochastic and there is no crisis. In the second type, the equilibrium price \( p \) is random and a crisis occurs with positive probability. Both types of equilibria will be shown to exist. Note that there cannot be equilibrium in which a crisis occurs for certain. If banks anticipated low asset prices with probability one, they would not be willing to hold long-term assets in the first place.

Another point worth noting is that this is a story about systemic or economy-wide crises, rather than individual bank runs. Each perfectly competitive bank takes the price \( p \) as given and, because we restrict attention to essential runs, the value of \( p \) determines whether the bank defaults. Only the failure of a non-negligible set of banks will affect the equilibrium price \( p \) in a way that makes individual defaults essential. From the point of view of the individual bank, the default is unavoidable and part of a general crisis. From the point of view of the banking sector as a whole, it is a shock without a cause.

So far, we have argued for the existence of multiple equilibria, only some of which are characterized by crises. We want to go further and suggest that some equilibria are more robust than others. To test for robustness we perturb the benchmark economy by introducing a small amount of exogenous aggregate uncertainty. More precisely, we assume the fraction of early consumers at a particular bank is a random variable \( \eta(\theta) = \alpha + \varepsilon \theta \), where \( \alpha \) is a bank-specific shock, \( \theta \) is an economy-wide shock, and \( \varepsilon \geq 0 \) is a constant. This model represents a family of economies with different amounts of aggregate, intrinsic uncertainty, depending on the value of \( \varepsilon \). We take as the benchmark the limit economy corresponding to \( \varepsilon = 0 \). In the limit economy, preference shocks are purely idiosyncratic. There is no aggregate intrinsic uncertainty. Economies corresponding to small positive values of \( \varepsilon \) are thought of as perturbations of the limit economy.

An equilibrium is **fundamental** if the endogenous variables are functions only of the primitives or fundamentals of the economy. An equilibrium that is not fundamental is referred to as a **sunspot equilibrium**.\(^5\) An equilibrium

\(^5\)Our definition implies that, in a sunspot equilibrium, the endogenous variables depend
of the limit economy is robust if a small perturbation of the economy leads to a small change in the equilibrium. More precisely, an equilibrium of the limit economy is robust if, for every $\delta > 0$ there exists $\epsilon_0 > 0$ such that for every value of $0 < \epsilon < \epsilon_0$ the $\epsilon$-perturbed economy has an equilibrium that is $\delta$-close to the equilibrium of the limit economy. The limit economy may have both fundamental and sunspot equilibria. There is no aggregate intrinsic uncertainty in the limit economy and, hence, no aggregate uncertainty in the fundamental equilibrium.\textsuperscript{6} In the absence of aggregate uncertainty, there can be no crises in a fundamental equilibrium.

To test the robustness of this equilibrium, we examine the equilibria of the perturbed economy for small values of $\epsilon$. We show that for any $\epsilon > 0$, however small, the equilibria of the perturbed economy always exhibit crises with positive probability. Furthermore, the probability of a crisis is bounded away from zero as $\epsilon$ becomes vanishingly small. Thus, in a robust equilibrium of the limit economy there must be extrinsic uncertainty and, under certain conditions, a non-negligible fraction of banks will default. The fundamental equilibrium of the limit economy is not robust.\textsuperscript{7}

To sum up, our theory combines attractive features from the two main strands of the literature. Like the sunspot approach, it produces large effects from small shocks. Like the real business cycle approach, it makes a firm prediction about the conditions under which crises will occur.

Our work is related to the wider literature on general equilibrium with incomplete markets (GEL). As is well known, sunspots do not matter when markets are complete. (For a precise statement, see Shell and Goenka, (1997)). The incompleteness in our model reveals itself in two ways. First, sunspots are assumed to be non-contractible, that is, the deposit contract is not explicitly contingent on the sunspot variable. In this respect we are simply

\textsuperscript{6}Here we have to be careful to distinguish the individual shocks from aggregate shocks. Each consumer observes a private preference shock and conditions his actions on that information. However, since an individual agent is negligible we can assume that aggregate variables are independent of individual shocks.

\textsuperscript{7}The limit economy has two types of equilibria with extrinsic uncertainty. In a trivial sunspot equilibrium, prices are random but the allocation is essentially the same as in the fundamental equilibrium and no banks default. In a non-trivial sunspot equilibrium, both prices and the allocation are random and default occurs with positive probability among a non-negligible set of banks.
following the incomplete contracts literature (see, for example, Hart (1995)). Secondly, there are no markets for Arrow securities contingent on the sunspot variable, so financial institutions cannot insure themselves against asset price fluctuations associated with the sunspot variable. This is the standard assumption of the GEI literature (see, for example, Geanakoplos et al. (1990) or Magill and Quinzii (1996)).

The rest of the paper is organized as follows. Section 2 contains the basic assumptions of the model. Section 3 describes the optimal contracts offered by banks and the rules governing default and liquidation. Section 4 defines equilibrium and Section 5 characterizes the equilibria of the limit economy, in which there is no aggregate uncertainty. The analysis of equilibrium in the perturbed economy is contained in Section 6. Here we show that in any equilibrium of the perturbed economy, crises occur with positive probability. We also show that the limit of a sequence of equilibria corresponding to a sequence of perturbed economies is an equilibrium in the limit economy and we characterize the limit equilibria. Section 7 contains concluding remarks. Proofs are gathered in Section 8.

2 Assets and preferences

The model we use is a special case of AG.

Dates. There are three dates $t = 0, 1, 2$ and a single good at each date. The good can be used for consumption or investment.

Assets. There are two assets, a short-term asset (the short asset) and a long-term asset (the long asset).

- The short asset is represented by a storage technology. Investment in the short asset can take place at date 1 or date 2. One unit of the good invested at date $t$ yields one unit at date $t + 1$, for $t = 0, 1$.

- The long asset takes two periods to mature and is more productive than the short asset. Investment in the long asset can only take place at date 0. One unit invested at date 0 produces $r > 1$ units at date 2.

Consumers. There is a continuum of ex ante identical consumers, whose measure is normalized to unity. Each consumer has an endowment $(1, 0, 0)$ consisting of one unit of the good at date 0 and nothing at subsequent dates.
There are two (ex post) types of consumers at date 1, *early consumers*, who only value consumption at date 1, and *late consumers*, who only value consumption at date 1. If \( \eta \) denotes the probability of being an early consumer and \( c_t \) denotes consumption at date \( t = 1, 2 \), the consumer’s ex ante utility is

\[
u(c_1, c_2, \eta) = \eta U(c_1) + (1 - \eta) U(c_2).
\]

The period utility function \( U : \mathbb{R}_+ \rightarrow \mathbb{R} \) is twice continuously differentiable and satisfies the usual neoclassical properties, \( U'(c) > 0, U''(c) < 0 \), and \( \lim_{c \searrow 0} U''(c) = \infty \).

**Uncertainty.** There are three sources of intrinsic uncertainty in the model. First, each individual consumer faces idiosyncratic uncertainty about his preference type (early or late consumer). Secondly, each bank faces idiosyncratic uncertainty about the number of early consumers among the bank’s depositors. Thirdly, there is aggregate uncertainty about the fraction of early consumers in the economy. Aggregate uncertainty is represented by a state of nature \( \theta \), a random variable with a continuous, positive density \( f(\theta) \) on the interval \([\theta_0, \theta_1]\), where \( \theta_0 < \theta_1 \) are finite numbers. The bank’s idiosyncratic shock is represented by a random variable \( \alpha \), with cumulative distribution function \( G(\alpha) \). The probability of being an early consumer at a bank in state \((\alpha, \theta)\) is denoted by \( \eta(\alpha, \theta) \), where

\[
\eta(\alpha, \theta) = \alpha + \varepsilon \theta,
\]

\( \varepsilon \geq 0 \) is a constant. We adopt the usual “law of large numbers” convention and assume that the fraction of early consumers at a bank in state \((\alpha, \theta)\) is identically equal to the probability \( \eta(\alpha, \theta) \). The economy-wide average of \( \alpha \) is assumed to be constant and equal to the mean \( \bar{\alpha} = \int_0^1 \alpha dG \). Thus, there is aggregate intrinsic uncertainty only if \( \varepsilon > 0 \).

**Information.** All uncertainty is resolved at date 1. The true value of \( \theta \) is publicly observed, the true value of \( \alpha \) for each bank is publicly observed, and each consumer learns his type, i.e., whether he is an early consumer or a late consumer.

### 3 Banking

Banks are financial institutions that provide investment and liquidity services to consumers. They do this by pooling the consumers’ resources, investing
them in a portfolio of short- and long-term assets, and offering consumers future consumption streams with a better combination of returns and liquidity than individual consumers could achieve by themselves. Banks also have access to the interbank capital market, from which consumers are excluded.\footnote{As Cone (1983) and Jacklin (1986) showed, if consumers have access to the capital market, it is impossible for banks to offer risk sharing that is superior to the market.}

Competition among banks forces them to maximize the welfare of their depositors. Anything a consumer can do, the bank can do. So there is no loss of generality in assuming that consumers deposit their entire endowment in a bank at date 0.\footnote{This is not simply an application of the Modigliani-Miller theorem. The consumer may do strictly better by putting all his “eggs” in the bank’s “basket”. Suppose that the deposit contract allows the individual to hold \(m\) units in the safe asset and deposit \(1 - m\) units in the bank. The bank invests \(y\) units in the short asset and \(1 - m - y\) in the long asset. If the bank does not default at date 1, the early consumers receive \(d + m\) and the late consumers receive \(p(\theta)r(1 - m - y) + y + \eta(\theta)d\). If the bank defaults, early and late consumers receive \(p(\theta)r(1 - m - y) + y + m\). Suppose that \(m > 0\) and consider a reduction in \(m\) and an increase in \(y\) and \(d\) of the same amount. It is clear that the early consumers’ consumption is unchanged. So is the late consumers’ consumption if the bank defaults. The change in the late consumers’ consumption when the bank does not default is

\[
\Delta y - \eta(\theta)\Delta d \quad \text{subject to} \quad \Delta m = -\frac{\Delta m}{p(\theta)} + \Delta m \geq 0
\]

because \(p(\theta) \leq 1\) and \(\Delta m < 0\). Thus, it is optimal for the bank to choose \(m = 0\).}

The bank invests \(y\) units per capita in the short asset and \(1 - y\) units per capita in the long asset and offers each consumer a deposit contract, which allows the consumer to withdraw either \(d_1\) units at date 1 or \(d_2\) units at date 2. Without loss of generality, we set \(d_2 = \infty\). This ensures that consumers receive the residue of the bank’s assets at date 2. Then the deposit contract is characterized by the payment \(d_1 = d\).

If \(p(\theta)\) denotes the price future consumption at date 1 in state \(\theta\), then the value of the bank’s assets at date 1 is \(y + p(\theta)r(1 - y)\).

A consumer’s type is private information. An early consumer cannot misrepresent his type because he needs to consume at date 1 but a late consumer can claim to be an early consumer, withdraw \(d\) at date 1 and store

\[
\frac{p(\theta)r(1 - m - y) + y - \eta(\theta)d}{(1 - \eta(\theta))p(\theta)} + m.
\]
it until date 2 when he can consume it. The deposit contract is incentive compatible if and only if the residual payment to late consumers at date 2 is at least \( d \). Incentive compatibility is consistent with the bank’s budget constraint if

\[
\eta(\alpha, \theta)d + (1 - \eta(\alpha, \theta))p(\theta)d \leq y + p(\theta)r(1 - y).
\]  

(1)

We assume that the late consumers withdraw at date 2 as long as it is incentive-compatible to do so, that is, as long as (1) is satisfied. If the constraint (1) cannot be satisfied, then clearly

\[ d > y + p(\theta)r(1 - y). \]

If all consumers demand \( d \) at date 1 the bank will be bankrupt. In the event of bankruptcy, the bank is assumed to liquidate its assets and divide them equally among the consumers. Each consumer receives \( y + p(\theta)r(1 - y) \).

Let \( x_t(d, y, \alpha, \theta) \) denote the consumption at date \( t \) if the bank chooses \((d, y)\) and the bank is in state \((\alpha, \theta)\). Let \( x = (x_1, x_2) \),

\[
x_1(d, y, \alpha, \theta) = \begin{cases} 
    d & \text{if (1) is satisfied} \\
    y + p(\theta)r(1 - y) & \text{otherwise}
\end{cases}
\]

and

\[
x_2(d, y, \alpha, \theta) = \begin{cases} 
    \frac{y + p(\theta)r(1 - y) - \eta d}{(1 - \eta)p(\theta)} & \text{if (1) is satisfied} \\
    y + p(\theta)r(1 - y) & \text{otherwise},
\end{cases}
\]

where \( \eta = \eta(\alpha, \theta) \). Using this notation, the bank’s decision problem (DP1) can be written as

\[
\max \ E [u(x(d, y, \alpha, \theta), \eta(\alpha, \theta))] \\
\text{s.t.} \quad 0 \leq d, 0 \leq y \leq 1.
\]

An ordered pair \((d, y)\) is optimal for the given price function \( p(\cdot) \) if it solves (DP1).

Implicit in our specification of the bank’s decision problem is the assumption that bank runs or crises are essential. To see this, note first that in equilibrium \( p(\theta) \leq 1 \); otherwise, banks could make an arbitrage profit by selling consumption forward and holding the short asset. As long as \( p(\theta) \leq 1 \), the cheapest way for the bank to make its promised payment at date 1 and satisfy incentive compatibility is to offer \( d \) to late consumers at date 2. This will satisfy the budget constraint if and only if (1) is satisfied.
4 Equilibrium

The bank’s decision problem is “non-convex”. To ensure the existence of equilibrium, we take advantage of the convexifying effect of large numbers and allow for possibility that ex ante identical banks will choose different deposit contracts \(d\) and portfolios \(y\). Each consumer is assumed to deal with a single bank and each bank offers a single contract. In equilibrium, consumers will be indifferent between banks offering different contracts. Consumers allocate themselves to different banks in proportions consistent with equilibrium.

To describe an equilibrium, we need some additional notation. A partition of consumers at date 0 is defined by an integer \(m < \infty\) and an array \(\rho = (\rho_1, ..., \rho_m)\) of numbers \(\rho_i \geq 0\) such that \(\sum_{i=1}^{m} \rho_i = 1\). Consumers are divided into \(m\) groups and each group \(i\) contains a measure \(\rho_i\) of consumers. We impose an arbitrary bound \(m\) on the number of groups to rule out pathological cases.\(^{10}\) The banks associated with group \(i\) offer a deposit contract \(d_i\) and a portfolio \(y_i\), both expressed in per capita terms. An allocation consists of a partition \((m, \rho)\) and an array \((d, y) = \{(d_i, y_i)\}_{i=1}^{m}\) such that \(d_i \geq 0\) and \(0 \leq y_i \leq 1\) for \(i = 1, ..., m\).

An allocation \((m, \rho, d, y)\) is attainable if it satisfies the market-clearing conditions

\[
\sum_i \rho_i E[\eta(\alpha, \theta)x_1(d_i, y_i, \alpha, \theta)] \leq \sum_i \rho_i y_i, \tag{2}
\]

and

\[
\sum_i \rho_i \{E[\eta(\alpha, \theta)x_1(d_i, y_i, \alpha, \theta) + (1 - \eta(\alpha, \theta)x_2(d_i, y_i, \alpha, \theta)]\}
\]

\[
= \sum_i \rho_i \{y_i + r(1 - y_i)\},
\]

for any state \(\theta\). In the market-clearing conditions, we take expectations with respect to \(\alpha\) because the cross-sectional distribution of idiosyncratic shocks is assumed to be the same as the probability distribution. The first inequality says that the total demand for consumption at date 1 is less than or equal to the supply of the short asset. The inequality may be strict, because an excess supply of liquidity can be re-invested in the short asset and consumed at date 2. The second inequality says that total consumption at date 2 is equal to the return from the investment in the long asset plus the amount

\(^{10}\)In general, two groups are sufficient for existence of equilibrium.
invested in the short asset at date 1, which is the difference between the left and right hand sides of (2).

We have already argued that \( p(\theta) \leq 1 \). If \( p(\theta) < 1 \) then no one is willing to invest in the short asset at date 1 and (2) must hold as an equation. A price function \( p(\cdot) \) is admissible (for the given allocation) if it satisfies the following complementary slackness condition:

For almost any state \( \theta, p(\theta) \leq 1 \) and \( p(\theta) < 1 \) implies that (2) holds as an equation.

Now we are ready to define an equilibrium.

An equilibrium consists of an attainable allocation \((m, \rho, d, y)\) and an admissible price function \( p(\cdot) \) such that, for every group \( i = 1, ..., m \), \((d_i, y_i)\) is optimal given the price function \( p(\cdot) \).

An equilibrium \((m, \rho, d, y, p)\) is pure if each group of banks makes the same choice:

\[
(d_i, y_i) = (d_j, y_j), \forall i, j = 1, ..., m.
\]

An equilibrium \((m, \rho, d, y, p)\) is semi-pure if the consumption allocations are the same for each group of banks:

\[
x(d_i, y_i, \alpha, \theta) = x(d_j, y_j, \alpha, \theta), \forall i, j = 1, ..., m.
\]

Otherwise, \((m, \rho, d, y, p)\) is a mixed equilibrium.

5 Equilibrium in the limit

In this section we characterize the equilibria of the limit economy in which \( \varepsilon = 0 \). There is no aggregate intrinsic uncertainty in the model, but there may still be aggregate extrinsic uncertainty (sunspots). We can classify equilibria in the limit economy according to the impact of extrinsic uncertainty. An equilibrium \((m, \rho, d, y, p)\) in the limit economy is a fundamental equilibrium (FE) if \( x(d_i, y_i, \alpha, \theta) \) is almost surely constant for each \( i \) and \( \alpha \) and if \( p(\theta) \) is almost surely constant. In that case, the sunspot variable \( \theta \) has no influence on the equilibrium values. An equilibrium \((m, \rho, d, y, p)\) of the limit economy is a trivial sunspot equilibrium (TSE) if \( x(d_i, y_i, \alpha, \theta) \) is almost surely constant for each \( i \) and \( \alpha \) and \( p(\theta) \) is not almost surely constant. In this case, the sunspot variable \( \theta \) has no effect on the allocation of consumption but it does affect the equilibrium price \( p(\theta) \). An equilibrium \((m, \rho, d, y, p)\) which is
neither a FE nor a TSE is called a non-trivial sunspot equilibrium (NTSE), that is, a NTSE is an equilibrium in which the sunspot variable has some non-trivial impact on the allocation of consumption.

We begin by considering the fundamental equilibrium in which, by definition, the price function \( p \) is almost surely constant: \( p(\theta) = \bar{p} \) for almost all \( \theta \). The FE referred to in the theorem is unique up to the choice of \( y_i \).

By definition, the price function \( p(\cdot) \) is almost surely equal to a constant \( \bar{p} \), say. In that case, it is easy to show that

\[
\bar{p} = \frac{1}{r}.
\]

A unit of the good invested in the long asset at date 0 will be worth \( \bar{pr} \) at date 1 and a unit of the good used to purchase the long asset at date 1 will be worth \( 1/\bar{pr} \) at date 2. If \( \bar{pr} = 1 = 1/\bar{pr} \), banks are indifferent between holding the two assets at date 0 and will only hold the long asset at date 1. If \( \bar{pr} > 1 \), banks will not hold the short asset at date 0; if \( \bar{pr} < 1 = 1/\bar{pr} \), banks will not hold the long asset at date 0 and will not hold the short asset at date 1. In either case, market clearing at date 1 is impossible.

Banks may choose \((d_i, y_i)\) so that they are forced to default in some states because of uncertainty about the idiosyncratic shock \( \alpha \). In order to distinguish crises caused by aggregate extrinsic uncertainty from defaults caused by idiosyncratic shocks, we will assume that the parameters are such that default is never optimal in the FE. In that case, the bank’s optimal choice of \((d_i, y_i)\) must satisfy the incentive constraint, so the bank’s optimal decision problem can be written as follows. At the equilibrium price \( \bar{p} = 1/r \), the value of the bank’s assets at date 1 is \( y_i + \bar{pr}(1 - y_i) = 1 \), independently of the choice of \( y_i \). In the absence of default, the budget constraint implies that the consumption at date 2 is given by \( r(1 - \alpha d_i)/(1 - \alpha) \). The incentive constraint requires that \( r(1 - \alpha d_i)/(1 - \alpha) \geq d_i \). Thus, the decision problem can written as:

\[
\max \quad E [\alpha U(d_i) + (1 - \alpha)U (r(1 - \alpha d_i)/(1 - \alpha))] \\
\text{st} \quad r(1 - \alpha d_i)/(1 - \alpha) \geq d_i, \forall \alpha.
\]

This is a convex programming problem and has a unique solution for \( d_i \). As noted, \( y_i \) is indeterminate, but the equilibrium allocation must satisfy the market-clearing condition

\[
\sum_i \rho_i E [\alpha d_i] = \sum_i \rho_i y_i.
\]
Note that there is a single pure FE \((\rho, m, d, y, \bar{p})\), in which \(y_i = E[\alpha d_i]\) for every \(i = 1, \ldots, m\).

### 5.1 Equilibrium without idiosyncratic shocks

In the special case where \(\alpha\) is a constant the following theorem partitions the equilibrium set into two cases with distinctive properties.

**Theorem 1** Suppose that \(\alpha\) is almost surely constant and let \((\rho, m, x, y, p)\) be an equilibrium of the limit economy, in which \(\varepsilon = 0\). There are two possibilities:

(i) \((\rho, m, x, y, p)\) is a semi-pure, fundamental equilibrium in which the probability of default is zero;

(ii) \((\rho, m, x, y, p)\) is a pure, trivial sunspot equilibrium in which the probability of default is zero;

**Proof.** See Section 8. □

By definition, an equilibrium must be either a FE, TSE, or NTSE. What Theorem 1 shows is that NTSE does not occur and each of the remaining cases is associated with distinctive properties in terms of symmetry and probability of default.

We have seen that the FE referred to in the theorem is unique up to the choice of \(y_i\). The banks in group \(i\) choose a non-stochastic consumption bundle \((c_1, c_2)\) to solve

\[
\begin{align*}
\max & \quad \alpha U(c_1) + (1 - \alpha)U(c_2) \\
\text{st} & \quad \alpha c_1 + (1 - \alpha)\bar{p}c_2 \leq 1 \\
& \quad c_1 \leq c_2.
\end{align*}
\]

The solution to this problem is uniquely determined by the first-order condition

\[
U'(c_1) - rU'(c_2) = 0,
\]

and the budget constraint

\[
\alpha c_1 + (1 - \alpha)\frac{c_2}{r} = 1.
\]

The first-order conditions imply that \(c_1 < c_2\). Thus, the incentive constraint is (strictly) satisfied and it is never optimal to allow for default.
The TSE is a variant of the unique pure FE. Let \( p(\cdot) \) be any non-constant price function such that \( p(\theta) < 1 \) almost surely and
\[
E \left[ \frac{1}{p(\theta)} \right] = r.
\]

Then \((\rho, m, x, y, p)\) is a TSE. Price fluctuations have no effect on the bank’s budget constraint, because there is no trade in the forward market at date 1. It is optimal to hold only the long asset between date 1 and date 2 because \( p(\theta) < 1 \) almost surely and the first-order conditions for a maximum with respect to \( y_i \) are satisfied. So \((x_i, y_i)\) is optimal for every group \( i \).

Whereas the value of \( y_i \) is indeterminate in FE, price variability requires every bank to choose the same value of \( y_i \) in the TSE. Otherwise, the budget constraint cannot be satisfied. Thus, a TSE is pure and unique.

Because the incentive constraint does not bind in either the FE or TSE, both achieve the first-best or Pareto-efficient allocation. No equilibrium can do better. Any bank can guarantee this level of utility by choosing \( \alpha d_i = y_i \), where \( d_i \) is the deposit contract chosen in the FE. For this choice of \((d_i, y_i)\) prices have no effect on the bank’s budget constraint and, because we assume crises are essential, the depositors will receive the first-best consumption. In a NTSE, by contract, agents receive noisy consumption allocations. Because they are risk averse, the noise in their consumption allocations is inefficient. Since we have seen that the bank can guarantee more to the depositors, this kind of equilibrium cannot exist.

### 5.2 Equilibrium with idiosyncratic shocks

In the case of idiosyncratic shocks, the following theorem partitions the equilibrium set into two cases, FE and NTSE.

**Theorem 2** Let \((\rho, m, d, y, p)\) be an equilibrium of the limit economy, in which \( \varepsilon = 0 \). There are three possibilities:

(i) \((\rho, m, d, y, p)\) is a semi-pure, fundamental equilibrium in which the probability of default is zero;

(ii) \((\rho, m, d, y, p)\) is a non-trivial sunspot equilibrium which is pure if the probability of default is zero.

**Proof.** See Section 8. ■
The fundamental equilibrium is semi-pure for the usual reasons and there is no default by assumption. Unlike the case with no idiosyncratic shocks, there can be no trivial sunspot equilibrium. To see this, suppose that some bank group with \( \rho_i \) chooses \((d_i, y_i)\) and that default occurs with positive probability. Then consumption for both early and late consumers is equal to \( y_i + p(\theta)r(1 - y_i) \), which is independent of \( p(\theta) \) only if \( y_i = 1 \). In that case there is no point in choosing \( d > 1 \) and so no need for default. If there is no probability of default, then the consumption of the late consumers is

\[
\frac{y_i + p(\theta)r(1 - y_i) - \alpha d_i}{(1 - \alpha)p(\theta)}.
\]

This is independent of \( p(\theta) \) only if \( y_i = \alpha d_i \), which cannot hold unless \( \alpha \) is a constant. Thus, there can be no TSE.

The only remaining possibility is a NTSE. If the probability of default is zero, then the bank’s decision problem is a convex programming problem and the usual methods suffice to show uniqueness of the optimum choice of \((d, y)\) under the maintained assumptions.

As long as there is no possibility of default, an equilibrium \((m, \rho, d, x, p)\) must be either a FE or a TSE. If there is a positive probability of default, the equilibrium \((m, \rho, d, x, p)\) must be a NTSE. Any NTSE is characterized by a finite number of prices. To see this, let \( B \subset \{1, \ldots, m\} \) denote the groups that default at some state \( \theta \). If \( p(\theta) < 1 \) then complementary slackness and the market-clearing condition (2) imply that

\[
\sum_{i \in B} \rho_i w_i(\theta) + \sum_{i \notin B} \alpha \rho_i d_i = \sum_i \rho_i y_i.
\]

Substituting for \( w_i(\theta) = y_i + p(\theta)r(1 - y_i) \) gives

\[
\sum_{i \in B} \rho_i (y_i + p(\theta)r(1 - y_i)) + \sum_{i \notin B} \alpha \rho_i d_i = \sum_i \rho_i y_i,
\]

or

\[
p(\theta) = \frac{\sum_i \rho_i y_i - \sum_{i \in B} \rho_i y_i - \sum_{i \notin B} \alpha \rho_i d_i}{\sum_{i \in B} \rho_i r(1 - y_i)},
\]

where the denominator must be positive. If not, then \( \sum_{i \in B} \rho_i r(1 - y_i) = 0 \), which implies that \( y_i = 1 \) for every \( i \in B \). The budget constraint implies that \( x_{i1}(\theta) \leq 1 \) almost surely, in which case there can be no default, contradicting
the definition of $B$. So the denominator must be positive. The expression on the right hand side can only take on a finite number of values because there is a finite number of sets $B$. Thus, there can be at most a finite number of different prices observed in equilibrium.

We have proved the following corollary.

**Corollary 3** Let $(\rho, m, d, y, p)$ be an equilibrium. There are three possibilities:

- In a FE, the range of $p(\cdot)$ is a singleton: $p(\theta) = 1/r$ almost surely;
- In a TSE, the range of $p(\cdot)$ is finite or infinite;
- In a NTSE, the range of $p(\cdot)$ is finite.

The properties of equilibria in the limit economy are summarized in the following table.

<table>
<thead>
<tr>
<th>Symmetry</th>
<th>Type</th>
<th>Default</th>
<th>Range of $p(\cdot)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semi-pure</td>
<td>FE</td>
<td>No</td>
<td>Singleton</td>
</tr>
<tr>
<td>Pure</td>
<td>TSE</td>
<td>No</td>
<td>Finite or infinite</td>
</tr>
<tr>
<td>Mixed or pure</td>
<td>NTSE</td>
<td>Possible</td>
<td>Finite</td>
</tr>
</tbody>
</table>

### 6 Limit of equilibria

In the special case $\varepsilon = 0$, there is idiosyncratic uncertainty about consumers’ types and the proportion of types at a given bank, but no aggregate uncertainty about the proportion of early consumers in the economy as a whole. When $\varepsilon > 0$ is small, the amount of aggregate uncertainty is positive but small. In this section, we test the robustness of an equilibrium in the limit by examining the limit of sequences of equilibria as $\varepsilon > 0$ becomes vanishingly small. If $\varepsilon > 0$ then, by definition, every equilibrium is fundamental because the state $\theta$ represents intrinsic uncertainty.

Introducing a small amount of uncertainty changes the analysis of equilibrium in one crucial respect: there must be default with positive probability. Suppose that $(m, \rho, d, y, p)$ is an equilibrium and suppose, contrary to what we want to prove, that default occurs with probability zero. On the one hand, the supply of liquidity at date 0 is fixed and independent of $\theta$: $\sum_i \rho_i y_i$. On the other hand, the demand for liquidity is inelastic and proportional to $\eta(\theta)$: $\sum_i \rho_i \eta(\theta) d_i$ where $d_i$ is the constant promised to early withdrawers in group $i$ at date 1. In equilibrium, the demand for liquidity must be less than
or equal to the supply at date 1, but since $\eta(\theta)$ is continuously distributed, this implies that there will be excess supply of liquidity with probability one. The excess liquidity will have to be re-invested in the short asset from date 1 until date 2. This is consistent with equilibrium only if the returns on the short and long assets are the same between date 1 and date 2, that is, $p(\theta) = 1$ with probability one. However, $p(\theta) = 1$ (almost surely) is inconsistent with equilibrium, because it implies that the long asset dominates the short asset between dates 0 and 1, no one holds the short asset at date 0 and there cannot be an excess supply of liquidity at date 1. This proves that in any equilibrium with $\varepsilon > 0$ there must be default with positive probability.

The next theorem characterizes the properties of equilibria in perturbed economies.

**Theorem 4** Let $(m, \rho, d, y, p)$ denote an equilibrium for a perturbed economy, in which $\varepsilon > 0$. Then $(m, \rho, d, y, p)$ is a mixed equilibrium in which the probability of default is strictly positive.

**Proof.** See Section 8. □

As a corollary, we can deduce several other properties of equilibrium with $\varepsilon > 0$. First, there cannot be a state in which all banks are in default, for this would imply $p(\theta) = 0$ which is inconsistent with equilibrium. Hence, any equilibrium must be mixed. Secondly, it will never be optimal to have default with probability one, so there must be a non-negligible set of states with no default in which, by the earlier argument, we have $p(\theta) = 1$. Finally, for any value of $\varepsilon > 0$, the volatility of asset prices, as measured by the variance of $p(\cdot)$, is bounded away from zero. Both assets are held at date 0 in equilibrium and this requires that the high returns to the long asset, associated with $p(\theta) = 1$ must be balanced by low returns associated with $p(\theta) < 1/r$. These properties are all preserved in the limit as $\varepsilon \to 0$.

The next result shows that the limit of a sequence of equilibria is an equilibrium in the limit.

**Theorem 5** Consider a sequence of perturbed economies corresponding to $\varepsilon = 1/q$, where $q$ is a positive integer, and let $(m^q, \rho^q, d^q, y^q, p^q)$ be an equilibrium for the corresponding perturbed economy. For some convergent sub-sequence $q \in Q$ let

$$(m^0, \rho^0, d^0, y^0, p^0) = \lim_{q \in Q}(m^q, \rho^q, d^q, y^q, p^q).$$
If \( \rho_0^i > 0, \, d_0^i > 0, \) and \( 0 < y_0^i < 1 \) for \( i = 1, \ldots, m^0 \) then \((m^0, \rho^0, x^0, y^0, p^0)\) is an equilibrium of the limit economy.

**Proof.** See Section 8. ■

The probability of default in each equilibrium \((m^q, \rho^q, d^q, y^q, p^q)\) is bounded away from both zero and one, so the same must be true in the limit equilibrium \((m^0, \rho^0, d^0, y^0, p^0)\). It follows immediately that \((m^0, \rho^0, d^0, y^0, p^0)\) is a NTSE of the limit economy.

### 7 Concluding remarks

In this paper, we have investigated the relationship between intrinsic and extrinsic uncertainty in a model of financial crises. Our general approach is to regard extrinsic uncertainty as a limiting case of intrinsic uncertainty. In our model, small shocks to the demand for liquidity are always associated with large fluctuations in asset prices. These price fluctuations cause financial crises to occur with positive probability. In the limit, as the liquidity shocks become vanishingly small, the model converges to one with extrinsic uncertainty. The limit economy has three kinds of equilibria,

- fundamental equilibria, in which there is neither aggregate uncertainty nor a positive probability of crisis;
- trivial sunspot equilibria, in which prices fluctuate but the real allocation is the same as in the fundamental equilibrium;
- and non-trivial sunspot equilibria, in which prices fluctuate and financial crises occur with positive probability.

Introducing small shocks into the limit economy destabilizes all but the third type of equilibrium. We argue that only the non-trivial sunspot equilibria are robust in the sense that a small perturbation of the model causes a small change in the equilibrium. This selection criterion provides an argument for the relevance of extrinsic uncertainty and the necessity of financial crises.

Although crises in the limit economy arise from extrinsic uncertainty, the causation is quite different from the bank run story of Diamond and Dybvig (1983). In the Diamond-Dybvig story, bank runs are spontaneous events that depend on the decisions of late consumers to withdraw early.
Given that almost all agents withdraw at date 1, early withdrawal is a best response for every agent; but if late consumers were to withdraw at date 2, then late withdrawal is a best response for every late consumer. So there are two “equilibria” of the coordination game played by agents at date 1, one with a bank run and one without. This kind of coordination failure plays no role in the present model. In fact, coordination failure is explicitly ruled out: a bank run occurs only if the bank cannot simultaneously satisfy its budget constraint and its incentive constraint. When bankruptcy does occur, it is the result of low asset prices. Asset prices are endogenous, of course, and there is a “self-fulfilling” element in the relationship between asset prices and crises. Banks are forced to default and liquidate assets because asset prices are low and asset prices are low as a result of mass bankruptcy and the association liquidation of bank assets.

One interesting difference between the present story and Diamond and Dybvig (1983) is that here a financial crisis is a systemic event. A crisis occurs only if the number of defaulting banks is large enough to affect the equilibrium asset price. In the Diamond-Dybvig model, by contrast, bank runs are an idiosyncratic phenomenon. Whether a run occurs at a particular bank depends on the decisions taken by the bank’s depositors. It is only by coincidence that runs are experienced by several banks at the same time.

At the heart of our theory is a pecuniary externality: when one group of banks defaults and liquidates its assets, it forces down the price of assets and this may cause another group of banks to default. This pecuniary externality may be interpreted as a form of contagion.

Allen and Gale (2000a) describes a model of contagion in a multi-region economy. Bankruptcy is assumed to be costly: long-term projects can be liquidated prematurely but a fraction of the returns are lost. This deadweight loss from liquidation creates a spillover effect in the adjacent regions where the claims on the bankrupt banks are held. If the spillover effect is large enough, the banks in the adjacent regions will also be forced into default and liquidation. Each successive wave of bankruptcies increases the loss of value and strengthens the impact of the spillover effect on the next region. Under certain conditions, a shock to one small region can propagate throughout the economy. By contrast, in the present model, a bank’s assets are always marked to market. Given the equilibrium asset price \( p \), bankruptcy does not change the value of the bank’s portfolio. However, if a group of banks defaults, the resulting change in the price \( p \) may cause other banks to default, which will cause further changes in \( p \), and so on. The “contagion” in both
models is instantaneous.

Several features of the model are special and deserve further consideration.

Pecuniary externalities “matter” in our model because markets are incomplete: if banks could trade Arrow securities contingent on the states $\theta$, they would be able to insure themselves against changes in asset values (Allen and Gale (2000d)). No trade in Arrow securities would take place in equilibrium, but the existence of the markets for Arrow securities would have an effect. The equilibrium allocation would be incentive-efficient, sunspots would have no real impact, and there would be no possibility of crises.

It is important that small shocks lead to large fluctuations in asset prices (and large pecuniary externalities). We have seen, in the case of trivial sunspot equilibria, that small price fluctuations have no real effect. What makes the pecuniary externality large in this example is inelasticity of the supply and demand for liquidity. Inelasticity arises from two assumptions. First, the supply of liquidity at date 1 is fixed by the decisions made at date 0. Secondly, the assumption of Diamond-Dybvig preferences implies that demand for consumption at date 1 is interest-inelastic. This raises a question about the robustness of the results when more general preferences are allowed.

One justification for the Diamond-Dybvig preferences is that they provide a cheap way of capturing, within the standard, Walrasian, auction-market framework, some realistic features of alternative market clearing mechanisms. In an auction market, prices and quantities adjust simultaneously in a *tatonnement* process until a full equilibrium is achieved. An alternative mechanism is one in which quantities are chosen before prices are allowed to adjust. An example is the use of market orders. If depositors were required to make a withdrawal decision before the asset price was determined in the interbank market, the same inelasticity of demand would be observed even if depositors had preferences that allowed for intertemporal substitution. There may be other institutional structures that have the qualitative features of our example. An investigation of these issues goes far beyond the scope of the present paper, but it is undoubtedly one of the most important topics for future research.
8 Proofs

8.1 Proof of Theorem 1

By definition, an equilibrium \((\rho, m, d, y, p)\) must be either a FE, TSE, or NTSE. The theorem is proved by considering each case in turn.

Case (i). If \((\rho, m, d, y, p)\) is a FE then by definition the price \(p(\theta)\) is almost surely constant and, for each group \(i\), the consumption allocation \(x(d_i, y_i, \alpha, \theta)\) is almost surely constant. In particular, \(x_1(d_i, y_i, \alpha, \theta) = d_i\) with probability one so there is no default in equilibrium.

Let \(\bar{p}\) denote the constant price and \(c_i = (c_{i1}, c_{i2})\) the consumption allocation chosen by banks in group \(i\). The decision problem of a bank in group \(i\) is

\[
\max \ E \left[ \alpha U(c_{i1}) + (1 - \alpha)U(c_{i2}) \right] \\
\text{s.t.} \quad c_{i1} \leq c_{i2}, 0 \leq y_i \leq 1 \\
\alpha c_{i1} + (1 - \alpha)\bar{p}c_{i2} \leq y_i + \bar{p}r(1 - y_i).
\]

Clearly, \(y_i\) will be chosen to maximize \(y_i + \bar{p}r(1 - y_i)\). Then the strict concavity of \(U(\cdot)\) implies that \(c_i\) is uniquely determined and independent of \(i\). Thus, the equilibrium is semi-pure: \(c_i = c_j\) for any \(i\) and \(j\).

Case (ii). Suppose that \((\rho, m, d, y, p)\) is a TSE. Then by definition \(p(\theta)\) is not almost surely constant and, for each group \(i\), the consumption allocation \(x(d_i, y_i, \alpha, \theta)\) is almost surely constant. In particular, \(x_1(d_i, y_i, \alpha, \theta) = d_i\) with probability one so there is no default in equilibrium.

Let \(c_i\) denote the consumption allocation chosen by banks in group \(i\). The budget constraint at date 1 reduces to

\[
\alpha c_{i1} - y_i = -p(\theta)((1 - \alpha)c_{i2} - r(1 - y_i)), \quad \text{a.s.}
\]

Since \(p(\theta)\) is not almost surely constant, this equation can be satisfied only if

\[
\alpha c_{i1} - y_i = (1 - \alpha)c_{i2} - r(1 - y_i) = 0.
\]

Then the choice of \((c_i, y_i)\) must solve the problem

\[
\max \ E \left[ \alpha U(c_{i1}) + (1 - \alpha)U(c_{i2}) \right] \\
\text{s.t.} \quad c_{i1} \leq c_{i2}, 0 \leq y_i \leq 1 \\
\alpha c_{i1} = y_i, (1 - \alpha)c_{i2} = r(1 - y_i).
\]

The strict concavity of \(U(\cdot)\) implies that this problem uniquely determines the value of \(c_i\) and hence \(y_i\), independently of \(i\). Consequently, the equilibrium is pure.

22
Case (iii). Suppose that \((\rho, m, d, y, p)\) is a NTSE. The allocation of consumption for group \(i\) is \(x(d_i, y_i, \alpha, \theta)\), and the expected utility of each group is the same

\[
E \left[ u(x(d_i, y_i, \alpha, \theta), \alpha) \right] = E \left[ u(x(d_j, y_j, \alpha, \theta), \alpha) \right], \forall i, j.
\]

The mean allocation \(\sum_i \rho_i x(d_i, y_i, \alpha, \theta)\) satisfies the market clearing conditions for every \(\theta\) and hence consumption bundle \(E \left[ \sum_i \rho_i x(d_i, y_i, \alpha, \theta) \right]\) is feasible for the planner. Since agents are strictly risk averse,

\[
u (E \left[ \sum_i \rho_i x(d_i, y_i, \alpha, \theta) \right], \alpha) > \sum_i \rho_i E \left[ u(x(d_i, y_i, \alpha, \theta), \alpha) \right].
\]

This contradicts the equilibrium conditions, since the individual bank could choose

\[
y_0 = \alpha d_0 = \alpha E \left[ \sum_i \rho_i x(d_i, y_i, \alpha, \theta) \right]
\]

and achieve a higher utility.

### 8.2 Proof of Theorem 2

Again we let \((\rho, m, d, y, p)\) be a fixed but arbitrary equilibrium and consider each of three cases in turn.

Case (i). If \((\rho, m, d, y, p)\) is a FE \(p(\theta)\) is almost surely constant and, for each group \(i\) and each \(\alpha\), the consumption allocation \(x(d_i, y_i, \alpha, \theta)\) is almost surely constant. In particular, \(x_1(d_i, y_i, \alpha, \theta) = d_i\) with probability one so there is no default in equilibrium.

Let \(\bar{p}\) denote the constant price and \(c_i(\alpha) = (c_{i1}(\alpha), c_{i2}(\alpha))\) the consumption allocation chosen by banks in group \(i\). The decision problem of a bank in group \(i\) is

\[
\begin{align*}
\max & \quad E \left[ \alpha U(c_{i1}(\alpha)) + (1 - \alpha)U(c_{i2}(\alpha)) \right] \\
\text{s.t.} & \quad c_{i1}(\alpha) \leq c_{i2}(\alpha), 0 \leq y_i \leq 1 \\
& \quad \alpha c_{i1}(\alpha) + (1 - \alpha)\bar{p}c_{i2}(\alpha) \leq y_i + \bar{p}(1 - y_i).
\end{align*}
\]

Clearly, \(y_i\) will be chosen to maximize \(y_i + \bar{p}(1 - y_i)\). Then the strict concavity of \(U(\cdot)\) implies that \(c_i(\alpha)\) is uniquely determined and independent of \(i\) (but not of \(\alpha\)). Thus, the equilibrium is semi-pure: \(c_i = c_j\) for any \(i\) and \(j\).
Case (ii). Suppose that \( (\rho, m, d, y, p) \) is a TSE. Then by definition \( p(\theta) \) is not almost surely constant and, for each group \( i \) and \( \alpha \), the consumption allocation \( x(d_i, y_i, \alpha, \theta) \) is almost surely constant. In particular, \( x_1(d_i, y_i, \alpha, \theta) = d_i \) with probability one so there is no default in equilibrium.

Let \( c_i(\alpha) \) denote the consumption allocation chosen by banks in group \( i \). The budget constraint at date 1 reduces to
\[
\alpha c_{i1}(\alpha) - y_i = -p(\theta)((1 - \alpha)c_{i2}(\alpha) - r(1 - y_i)), \text{ a.s.}
\]
Since \( p(\theta) \) is not almost surely constant, this equation can be satisfied only if
\[
\alpha c_{i1}(\alpha) - y_i = (1 - \alpha)c_{i2}(\alpha) - r(1 - y_i) = 0.
\]
Since \( \alpha \) is not constant this can only be true if \( c_{i1}(\alpha) = c_{i2}(\alpha) = 0 \), a contradiction. Thus, there cannot be a TSE when \( \alpha \) is not constant.

Case (iii). The only remaining possibility is that \( (\rho, m, d, y, p) \) is a NTSE. If there is no default in this equilibrium, then each bank in group \( i \) solves the problem
\[
\max_{d_i, y_i} E\left[ \alpha U(d_i) + (1 - \alpha)U\left(\frac{y_i + p(\theta)(1 - y_i) - \alpha d_i}{(1 - \alpha)p(\theta)}\right)\right]
\]
subject to
\[
\frac{y_i + p(\theta)(1 - y_i)}{(1 - \alpha)p(\theta)} \geq d_i.
\]
This is a convex programming problem and it is easy to show that the strict concavity of \( U(\cdot) \) uniquely determines \((d_i, y_i)\). Thus a NTSE without default is pure.

8.3 Proof of Theorem 4

The proof is by contradiction. Suppose, contrary to what we want to prove, that the probability of default in \((\rho, m, d, y, p)\) is zero. Then for each group \( i \), \( x_1(d_i, y_i, \alpha, \theta) = d_i \) almost surely and the market-clearing condition (2) implies
\[
\sum_i E[p_i \eta(\alpha, \theta)] d_i \leq \sum_i p_i y_i. \tag{3}
\]
There are two cases to consider. In the first case, \( \sum_i p_i y_i = 0 \). Then \( d_i = 0 \) for every \( i \) and the utility achieved in equilibrium is
\[
E \left[ \eta(\alpha, \theta)U(0) + (1 - \eta(\alpha, \theta))U\left(\frac{y}{1 - \eta(\alpha, \theta)}\right)\right].
\]
By holding a small amount $\delta > 0$ of the short asset, positive consumption could be guaranteed at the first date. Optimality requires that

$$E \left[ \eta(\alpha, \theta) U \left( \frac{\delta}{\eta(\alpha, \theta)} \right) \right] + E \left[ (1 - \eta(\alpha, \theta)) U \left( \frac{r(1 - \delta)}{1 - \eta(\alpha, \theta)} \right) \right]$$

$$\leq E \left[ \eta(\alpha, \theta) U(0) + (1 - \eta(\alpha, \theta)) U \left( \frac{r}{1 - \eta(\alpha, \theta)} \right) \right].$$

for any $\delta > 0$. In the limit as $\delta \to 0$,

$$E [U'(0)] - E \left[ U' \left( \frac{r}{1 - \eta(\alpha, \theta)} \right) \right] \leq 0,$$

which contradicts the assumption that $U'(0) = \infty$.

In the second case, $\sum_i \rho_i y_i > 0$. Then the market-clearing condition (3) and the fact that $\eta(\alpha, \theta) = \alpha + \varepsilon \theta$ together imply that

$$\sum_i \rho_i \eta(\alpha, \theta)d_i < \sum_i \rho_i y_i$$

with probability one. The complementary slackness condition implies that $p(\theta) = 1$ almost surely, in which case the long asset dominates the short asset at date 0. Then banks will only hold the long asset at date 0 and the goods market cannot clear at date 1, a contradiction.

Thus, the probability of default must be positive.

8.4 Proof of Theorem 5

Continuity and the convergence of $\{(\rho_i^0, m, d_i^0, y_i^0, p_i^0)\}$ immediately implies the following properties of the limit point $(\rho^0, m, d^0, y^0, p^0)$:

(i) $\sum_i \rho_i^0 = 1$ and $\rho_i^0 \geq 0$ for every $i$ so $(\rho^0, m)$ is a partition.

(ii) For every $i$, $(d_i^0, y_i^0) \in \mathbb{R}_+ \times [0, 1]$. The market-clearing conditions

$$\sum_i \rho_i^0 E \left[ \alpha x_1(d_i^0, y_i^0, \alpha, \theta) | \theta \right] \leq \sum_i \rho_i^0 y_i^0,$$

and

$$\sum_i \rho_i^0 E \left\{ \alpha x_1(d_i^0, y_i^0, \alpha, \theta) + (1 - \alpha)x_2(d_i^0, y_i^0, \alpha, \theta) | \theta \right\} = \sum_i \rho_i^0 \{y_i^0 + r(1 - y_i^0)\}$$

25
are satisfied in the limit and the complementary slackness condition holds. Thus, \((d^0, y^0)\) is an attainable allocation.

It remains to show that \((d^0_i, y^0_i)\) is optimal for each \(i\). Let \(W^q(d_i, y_i, \theta)\) denote the utility associated with the pair \((d_i, y_i)\) in the perturbed economy corresponding to \(\varepsilon = 1/q\) when the price function is \(p^q\) and let \(W^0(d_i, y_i, \theta)\) denote the utility associated with the pair \((d_i, y_i)\) in the limit economy corresponding to \(\varepsilon = 0\) when the price function is \(p^0\). The function \(W^q(\cdot)\) is discontinuous at the bankruptcy point defined implicitly by the condition

\[
(\alpha + (1 - \alpha)p^0(\theta)) d_i = y_i + p^0(\theta)r(1 - y_i). \tag{4}
\]

If (4) occurs with probability zero in the limit, then it is easy to see from the assumed convergence properties that

\[
W^q(d_i, y_i, \theta) \to W^0(d_i, y_i, \theta), \text{ a.s.}
\]

and hence

\[
\lim_{q \to \infty} E[W^q(d_i, y_i, \theta)] = E[W^0(d_i, y_i, \theta)].
\]

Let \((d^0_i, y^0_i)\) denote the pair corresponding to the limiting consumption allocation \(x^0_i\) and let \(\{(d^q_i, y^q_i)\}\) denote the sequence of equilibrium choices converging to \((d^0_i, y^0_i)\). There may exist a set of \(\theta\) with positive measure such that \((d^q_i, y^q_i)\) implies default in state \(\theta\) for arbitrarily large \(q\) but that \((d^0_i, y^0_i)\) does not imply default in state \(\theta\). Then at least we can say that

\[
\liminf_{q \to \infty} W^q(d^q_i, y^q_i, \theta) \leq W^0(d^0_i, y^0_i, \theta), \text{ a.s.}
\]

and this implies that

\[
\liminf_{q \to \infty} E[W^q(d^q_i, y^q_i, \theta)] \leq E[W^0(d^0_i, y^0_i, \theta)].
\]

Now suppose, contrary to what we want to prove, that \((d^0_i, y^0_i)\) is not optimal. Then there exists a pair \((d_i, y_i)\) such that \(E[W^0(d_i, y_i, \theta)] > E[W^0(d^0_i, y^0_i, \theta)]\). If \(d_i = 0\) then (4) is satisfied and it is clear that for some sufficiently large value of \(q\), \(E[W^0(d_i, y_i, \theta)] > E[W^0(d^0_i, y^0_i, \theta)]\), contradicting the equilibrium conditions. If \(d_i > 0\), then either the critical condition (4) is satisfied or we can find a slightly lower value \(d' < d\) that does satisfy the critical condition. To see this, note first that the critical condition uniquely determines the value of \(p^0(\theta)\) as long as

\[
(1 - \alpha)d \neq r(1 - y)
\]
which is true for almost every value of \( d_i \). Secondly, if the value of \( p^0(\theta) \) for which the critical condition is satisfied is an atom, we can always find a slightly smaller value \( d'_i < d_i \) such that the value of \( p^0(\theta) \) for which the critical condition is satisfied is not an atom. Furthermore, reducing \( d \) slightly will at most reduce the payoff by a small amount, so for \( d'_i < d_i \) and close enough to \( d_i \) we still have \( E[W^0(d'_i, y_i, \theta)] > E[W^0(d_i, y_i, \theta)] \). Then this leads to a contradiction in the usual way.

References


