Banking and Markets

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

March 27, 2001

Abstract

This paper integrates a number of recent themes in the literature on banking and asset markets—optimal risk sharing, limited market participation, asset-price volatility, market liquidity, and financial crises—in a general-equilibrium theory of the financial system. A complex financial system comprises both financial markets financial institutions. Financial institutions can take the form of intermediaries or banks. Banks, unlike intermediaries, are subject to runs, but crises do not imply market failure. We show that a sophisticated financial system—a system with complete markets for aggregate risk and limited market participation—is incentive-efficient, if the institutions take the form of intermediaries, or else constrained-efficient, if they take the form of banks. We also consider an economy in which the markets for aggregate risks are incomplete. In this context, there is a role for prudential regulation: regulating liquidity can improve welfare.

1 Markets, intermediaries and crises

Financial crises in Southeast Asia, Russia, and elsewhere have revived interest in the theory of financial crises. To understand the origins of financial crises, we need to study complex, decentralized, financial systems comprising both financial markets and financial institutions.[1] For the most part, first-generation models of financial crises, beginning with Bryant (1980) and Diamond and Dybvig (1983), analyze the behavior of a single representative bank, are partial equilibrium in nature, and consist of a contracting problem followed by a coordination problem.[2] We have two objectives in this paper:

- to develop a model of a complex financial system in which there is a role for both financial institutions and financial markets;
- and to provide a microeconomic welfare analysis of financial crises and optimal stabilization policy.

We combine recent developments in the theory of banking with further innovations to model a complex financial system. The model has several interesting features.

[1] In this paper we use the term "financial markets" narrowly to denote markets for securities. Other authors have allowed for markets in which mechanisms are traded (e.g., Bisin and Gottardi (2000)). We prefer to call this intermediation. Formally, the two activities are similar, but in practice the economic institutions are quite different.

[2] First generation models of financial crises were developed in the 1980s, beginning with seminal work on bank runs by Bryant (1980) and Diamond and Dybvig (1983). Important contributions were also made by Chari and Jagannathan (1988), Chari (1989), Champ, Smith, and Williamson (1996), Jacklin (1986), Jacklin and Bhattacharya (1988), Postlewaite and Vives (1986), Wallace (1988; 1990) and others. Theoretical research on speculative currency attacks, banking panics, the role of liquidity and contagion have taken a number of approaches. One is built on the foundations provided by early research on bank runs (e.g., Diamond (1997), Allen and Gale (1998; 1999; 2000a; 2000b), Chang and Velasco (1998a; 1998b)), Diamond and Rajan (2000) and Peck and Shell (1999)). Other approaches include those based on macroeconomic models of currency crises that developed from the insights of Krugman (1979), Obstfeld (1986) and Calvo (1988) (see, e.g., Corsetti, Pesenti, and Roubini (1999) for a recent contribution and Flood and Marion (1999) for a survey), game theoretic models (see Morris and Shin (1998), Morris (2000) and Morris and Shin (2000) for an overview), amplification mechanisms (e.g., Cole and Kehoe (1995) and Chari and Kehoe (2000)) and the borrowing of foreign currency by firms (e.g., Aghion, Bacchetta and Banerjee (2000)).

- It introduces institutions into a general-equilibrium theory of markets. Just as general-equilibrium theory tends to ignore financial institutions, the theory of banking largely ignores markets. In order to understand the operation of complex, decentralized, financial systems we need to integrate the theory of banking with the theory of asset markets.

- It endogenizes the cost of liquidation. Most of the literature, following Diamond and Dybvig (1983), assumes the existence of a technology for liquidating projects. Here we assume that a financially distressed institution sells assets to other institutions. This realistic feature of the model has important implications for welfare analysis. Ex post, liquidation does not entail a deadweight cost because assets are merely transferred from one owner to another. Ex ante, liquidation can result in inefficient risk sharing, but only if markets for hedging the risk are incomplete.

- It allows for a more general specification of the economic environment, thus allowing us to test the robustness of results obtained using the more specialized models of the banking literature. In particular, it allows us to test robustness to the introduction of financial markets.

- It allows for interaction between liquidity and asset pricing. In particular, there is a role for cash-in-the-market pricing (Allen and Gale, 1994). The asset market is comprised of financial institutions and their portfolio choices collectively determine how much liquidity is available in the market. For the individual institution, however, the market is a source of liquidity, because the institution can liquidate long-term assets by selling them to the market.

- It allows us to analyze the regulation of the financial system using the standard tools of welfare economics. Instead of asking how the central bank can achieve stability or control interest rates, which indirectly affect welfare, we directly analyze how the policy achieves microeconomic efficiency or impacts social welfare.

- It has important implications for the positive analysis of the financial system, e.g., the role of "mixed" equilibria in which identical banks choose very different risk strategies.

For a long time economic thinking on crises has taken it as axiomatic that crises are best avoided. Recently, there has been some re-thinking of this axiom. This new thinking may be compared to the "fail safe" principle in automobile design. Since we know that accidents will happen, we design cars so that they crumple evenly, preventing secondary collisions that are often more damaging to the occupants than the initial collision. Wouldn't it be better to design our financial systems to be "fail safe", so that if an "accident" happens it doesn't cause unnecessary havoc? We have argued in a number of places that the welfare cost of financial crises is associated with inefficient liquidation of assets and suboptimal risk sharing, and not with crises per se (e.g., Allen and Gale, 1998). Hellwig (1994) makes a similar point when arguing that non-contingent deposit contracts misallocate risk across different classes of depositors.

Another cost of government intervention is the distortion it causes to the normal functioning of the financial system. Financial institutions and financial markets exist to facilitate the efficient allocation of risks and resources. A policy that aims to prevent financial crises has an impact on the normal functioning of the financial system. One of the advantages of a microeconomic analysis of financial crises is that it clarifies the costs associated with these distortions.

More generally, we argue that positive analysis (what causes crises? what happens when a crisis occurs?) needs to be complemented by normative analysis (what is the optimal policy towards crises? what is the optimal institutional structure when crises can occur?). Macroeconomic analyses have often focused on mistakes in government policy. A microeconomic approach emphasizes how the financial system works and why and presents crises as one aspect of a general account of financial activity. Welfare-oriented questions about the optimal design of financial systems and the nature of an optimal policy are naturally posed in a microeconomic framework.

Following Diamond and Dybvig (1983), we model financial institutions as providers of liquidity and risk sharing. We distinguish between *intermediaries*, which can offer general, incentive-compatible, risk-sharing contracts, and *banks*, which can only offer demand deposits. Banks are subject to runs, whereas intermediaries are not. However, this is not a source of market failure. We can show that a sophisticated financial system provides optimal liquidity and risk sharing. A financial system is "sophisticated" if markets for aggregate risks are complete and market participation is incomplete. Efficiency does not depend on whether financial institutions are intermediaries or banks. So there is a role for crises even in an efficient world.

Bordo (2000) observes that, historically, financial crises have caused extensive disruption to the economies in which they occurred. The same is true today in emerging markets. However, advanced economies have recently experienced crises that caused far less disruption. Why is this? One explanation is the growing sophistication of the financial systems in advanced economies. Financial markets allow institutions to implement sophisticated risk-management strategies that replicate the effect of general, incentive-compatible, risk-sharing contracts and thus avoid the costs of the non-contingent contracts that characterized the financial system in the past.

Historical evidence suggests that central banks have been important in controlling financial crises. This raises the question: What can the central bank or the government do that private institutions and the market cannot do? Our efficiency theorem assumes that markets for aggregate risk are complete. Missing markets may provide a role for government intervention. So, one interpretation of the efficiency theorem is that central banks serve as a substitute for missing markets. In that case, Bordo's observation can be explained by the increasing sophistication of the techniques of central banking.

One distinctive feature of our approach is that we focus on essential crises. Following Allen and Gale (1998), we call a crisis *essential* if, for certain parameter values, every equilibrium of the model is characterized by a crisis. Restricting attention to situations in which crises are essential gives the theory greater predictive power. Another view, following Bryant (1980) and Diamond and Dybvig (1983), emphasizes the existence of multiple equilibria. For a given set of parameter values, there exist equilibria with and without crises. One view is that equilibrium selection is a matter of market "psychology". Morris and Shin (1998) have argued to the contrary that asymmetric information leads to a unique equilibrium selection. In either case, weak fundamentals alone are not sufficient for a crisis; but, in the presence of weak fundamentals, crises can be triggered by self-fulfilling expectations.

The rest of the paper is organized as follows. Section 2 describes the primitives of the model. Section 3 explores the welfare properties of liquidity provision and risk sharing in the context of an economy with a sophisticated financial system in which institutions take the form of general intermediaries. Section 4 extends this analysis to an economy with a sophisticated financial system in which institutions take the form of banks. Section 5 shows that incomplete participation is critical for the optimality results achieved in the previous two sections. In Section 6 we analyze a simplified version of the economy with incomplete markets. We characterize conditions under which laisser-faire provision of liquidity is inadequate and show that government intervention is potentially Pareto-improving. Finally, in Section 7, we discuss a number of examples to illustrate and develop interesting aspects of the theory. Some of the proofs are gathered together in Section **??**.

2 The basic economy

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

The economy is subject to two kinds of uncertainty. First, individual agents are subject to idiosyncratic preference shocks, which affects their demand for liquidity (these will be described later). Second, the entire economy is subject to aggregate shocks that affect asset returns and the cross-sectional distribution of preferences. The aggregate shocks are represented by a finite number of states of nature, indexed by $\eta \in H$. At date 0, all agents have a common prior probability density $\nu(\eta)$ over the states of nature. All uncertainty is resolved at the beginning of date 1, when the state η is revealed and each agent discovers his individual preference shock.

Each agent has an endowment of one unit of the good at date 0 and no endowment at dates 1 and 2. So, in order to provide consumption at dates1 and 2, they need to invest.

There are two assets distinguished by their returns and liquidity structure. One is a short-term asset (the *short asset*), and the other is a long-term asset (the *long asset*). The short asset is represented by a storage technology: one unit invested in the short asset at date $t = 0, 1$ yields a return of one unit at date $t + 1$. The long asset yields a return after two periods. One unit of the good invested in the long asset at date 0 yields a random return of $R(\eta) > 1$ units of the good at date 2 if state η is realized.

Investors' preferences are distinguished ex ante and ex post. At date 0 there is a finite number n of types of investors, indexed by $i = 1, ..., n$. We call i an investor's *ex ante* type. An investor's ex ante type is common knowledge and hence contractible.. The measure of investors of type i is denoted by $\mu_i > 0$. The total measure of investors is normalized to one so that $\sum_i \mu_i = 1$.

While investors of a given ex ante type are identical at date 0, they receive a private, idiosyncratic, preference shock at the beginning of date 1. The date 1 preference shock is denoted by $\theta_i \in \Theta_i$, where Θ_i is a finite set. We call θ_i the investor's *ex post* type. Because θ_i is private information, contracts cannot be explicitly contingent on θ_i.

Investors only value consumption at dates 1 and 2. An investor's preferences are represented by a von Neumann-Morgenstern utility function $u_i(c_1, c_2; \theta_i)$, where c_t denotes consumption at date $t = 1, 2$. The utility function $u_i(\cdot; \theta_i)$ is assumed to be concave, increasing, and continuous for every type θ_i. Diamond and Dybvig (1983) used the idea of a preference shock θ_i to model liquidity preference.

The probability of being an investor of type (i, θ_i) conditional on state η is denoted by $\lambda_i(\theta_i, \eta) > 0$. The probability of being an agent of type i is μ_i. Consistency therefore requires that

$$\sum_{\theta_i} \lambda_i(\theta_i, \eta) = \mu_i, \forall \eta \in H.$$

By the usual "law of large numbers" convention, the cross-sectional distribution of types is assumed to be the same as the probability distribution λ. We can therefore interpret $\lambda_i(\theta_i, \eta)$ as the number of agents of type (i, θ_i) in state η.

3 Optimal intermediation

In this section we assume that financial institutions take the form of general intermediaries. Each intermediary offers a single contract and each ex ante type is attracted to a different intermediary.

One can, of course, imagine a world in which a single "universal" intermediary offer contracts to all ex ante types of investors. A universal intermediary could act as a central planner and implement the first-best allocation of risk. There would be no reason to resort to markets at all. Our world view is based on the assumption that transaction costs preclude this kind of centralized solution and that decentralized intermediaries are restricted in the number of different contracts they can offer. This assumption provides a role for financial markets in which financial intermediaries can share risk and obtain liquidity.

At the same time, financial markets alone will not suffice to achieve optimal risk sharing. Because individual economic agents have private information, markets for individual risks are incomplete. The markets that are available will not achieve an incentive-efficient allocation of risk. Intermediaries, by contrast, can offer individuals incentive-compatible contracts and improve on the risk sharing provided by the market.

In the Diamond-Dybvig (1983) model, all investors are ex ante identical. Consequently, a single representative bank can provide complete risk sharing and there is no need for markets to provide cross-sectional risk sharing across banks.

Allen and Gale (1994) showed that differences in risk and liquidity preferences can be crucial in explaining the behavior of asset markets. This is another reason for allowing for ex ante heterogeneity.

3.1 Markets

At date 0 investors deposit their endowments with an intermediary in exchange for a general risk sharing contract. The intermediaries have access to a complete set of Arrow securities markets at date 0. For each aggregate state η there is a security traded at date 0 that promises one unit of the good at date 1 if state η is observed and nothing otherwise. Let $q(\eta)$ denote the price of one unit of the Arrow security corresponding to state η, that is, the number of units of the good at date 0 needed to buy one unit of the good in state η at date 1.

All uncertainty is resolved at the beginning of date 1. Consequently there is no need to trade contingent securities at date 1. Instead, we assume there is a spot market and a forward market for the good at date 1. The good at date 1 is the numeraire so $p_1(\eta) = 1$; the price of the good at date 2 for sale at date 1 is denoted by $p_2(\eta)$, i.e., $p_2(\eta)$ is the number of units of the good at date 1 needed to purchase one unit of the good at date 2 in state η. Let $p(\eta) = (p_1(\eta), p_2(\eta)) = (1, p_2(\eta))$ denote the vector of goods prices at date 1 in state η.

Note that we do not assume the existence of a technology for physically liquidating projects. Instead, we follow Allen and Gale (1998) in assuming that an institution in distress sells long-term assets to other institutions. From the point of view of the economy as a whole, the long-term assets cannot be liquidated—some one has to hold them.

3.2 Intermediation mechanisms

Investors participate in markets indirectly, through intermediaries. An intermediary is a risk-sharing institution that invests in the short and long assets on behalf of investors and provides them with consumption at dates 1 and 2. Intermediaries use markets to hedge the risks that they manage for investors.

Each investor of type i gives his endowment (one unit of the good) to an intermediary of type i at date 0. In exchange, he gets a bundle of goods $x_i(\theta_i, \eta) \in \mathbf{R}^2_+$ at dates 1 and 2 in state η if he reports the ex post type θ_i. In effect, the function $x_i = \{x_i(\theta_i, \eta)\}$ is a *direct mechanism* that maps agents' reports into feasible consumption allocations.[3]

A feasible mechanism is incentive-compatible. The appropriate definition of the incentive-compatibility constraint depends on whether the agents can use the short asset to store the good from the first date to the second. If not, then we can assume that an agent who reports $\hat{\theta}_i$ in state η will consume $x_i(\hat{\theta}_i, \eta)$. Incentive compatibility simply says that an agent does not gain from misrepresenting his type:

$$u_i(x_i(\theta_i, \eta), \theta_i) \geq u_i(x_i(\hat{\theta}_i, \eta), \theta_i), \forall \theta_i, \hat{\theta}_i \in \Theta_i, \forall \eta \in H. \tag{1}$$

In other words, it is optimal for an investor to report his ex post type θ_i truthfully in each state η at date 1. Let X_i denote the set of incentive-compatible mechanisms when the storage technology is not available to agents.

Alternatively, we can assume that agents can use the short asset to save the good from date 1 to date 2. Suppose that the agent receives a consumption bundle $x_i(\theta_i, \eta)$ from the intermediary. By saving, he can obtain any consumption bundle $c \in \mathbf{R}^2_+$ such that

$$c_1 \leq x_{i1}(\theta_i, \eta), \sum_{t=1}^{2} c_t \leq \sum_{t=1}^{2} x_{it}(\theta_i, \eta). \tag{2}$$

Let $C(x_i(\theta_i, \eta))$ denote the set of consumption bundles satisfying (2). The maximum utility that can be obtained from the consumption bundle $x_i(\theta_i, \eta)$

[3] A director mechanism is normally a function that assigns a unique outcome to each profile of types chosen by the investors. In a symmetric direct mechanism, the outcome for a single investor depends only on the individual's report and the distribution of reports by other investors. In a truth-telling equilibrium, the reports of other investors is given by the distribution $\lambda(\cdot, \eta)$ so a symmetric direct mechanism should properly be written $x_i(\theta_i, \lambda(\cdot, \eta), \eta)$ but since $\lambda(\cdot, \eta)$ is given as a function of η there is no loss of generality in suppressing the reference to $\lambda(\cdot, \eta)$.

by saving is denoted by $u_i^*(x_i(\theta_i,\eta),\theta_i)$ and defined by

$$u_i^*(x_i(\theta_i,\eta),\theta_i) = \sup\{u_i(c,\theta_i) : c \in C(x_i(\theta_i,\eta))\}.$$

Then the incentive constraint with saving can be written by replacing $u_i(\cdot)$ with $u_i^*(\cdot)$ in (1):

$$u_i(x_i(\theta_i,\eta),\theta_i) \geq u_i^*(x_i(\hat{\theta}_i,\eta),\theta_i), \forall\theta_i,\hat{\theta}_i \in \Theta_i, \forall\eta \in H. \quad (3)$$

Notice that by placing $u_i(\cdot)$ on the left hand side of (3), we ensure that even a truth-telling agent will not want to save outside of the intermediary. There is no loss of generality in this restriction, since the intermediary can tailor the timing of consumption to the agent's needs. Let X_i^* denote the set of incentive-compatible mechanisms when the depositor has access to the storage technology. In what follows we state results for the case where depositors do not have access to the storage technology, but all of these results remain valid if X_i is replaced by X_i^*.

3.3 Equilibrium

Competition and free entry force the intermediaries to maximize the welfare of the typical depositor.

Recall that each intermediary issues a single contract and serves a single ex ante type of investor. We denote by i the representative intermediary that trades with the ex ante investor type i. The representative intermediary i takes in deposits of μ_i units of the good at date 0 and invests in $y_i \geq 0$ units of the short asset and $\mu_i - y_i \geq 0$ units of the long asset. In exchange it offers depositors an incentive-compatible mechanism x_i. Given the prevailing prices (p, q), the choice of mechanism x_i and investment portfolio y_i must satisfy the budget constraint:

$$\sum_\eta q(\eta) \sum_{\theta_i} \lambda(\theta_i,\eta)p(\eta)\cdot x_i(\theta_i,\eta) \leq \sum_\eta q(\eta)p(\eta)\cdot(y_i,(\mu_i - y_i)R(\eta)). \quad (4)$$

In state η, the cost of goods given to investors who report θ_i is $p(\eta)\cdot x_i(\theta_i,\eta)$ and there are $\lambda(\theta_i,\eta)$ such agents, so summing across ex post types θ_i we get the total cost of the mechanism in state η as $\sum_{\theta_i}\lambda(\theta_i,\eta)p(\eta)\cdot x_i(\theta_i,\eta)$. Multiplying by the cost of one unit of the good at date 1 in state η and summing over states η gives the total cost of the mechanism, in terms of units of the good at date 0, as the left hand side of equation (4). The right hand side is the total value of investments by the intermediary. In state η the short asset yields y_i units of the good at date 1 and the long asset yields $(\mu_i - y_i)R(\eta)$ units of the good at date 2 so the total value of the portfolio is $p(\eta) \cdot (y_i, (\mu_i - y_i)R(\eta))$. Multiplying by the price of a unit of the good at date 1 in state η and summing across states gives the total value of the investments by the intermediary, in terms of units of the good at date 0.

An *intermediated allocation* specifies an incentive-compatible mechanism x_i and a feasible portfolio y_i for each representative intermediary $i = 1, ..., n$. An intermediated allocation $\{(x_i, y_i)\}$ is *attainable* if it satisfies the market-clearing conditions at dates 1 and 2, that is, for each intermediary i,

$$\sum_i \sum_{\theta_i} \lambda_i(\theta_i, \eta) x_{i1}(\theta_i, \eta) \leq \sum_i y_i, \forall \eta \tag{5}$$

and

$$\sum_i \sum_{\theta_i} \lambda_i(\theta_i, \eta)(x_{i1}(\theta_i, \eta) + x_{i2}(\theta_i, \eta)) = \sum_i y_i + (\mu_i - y_i)R(\eta), \forall \eta. \tag{6}$$

Condition (5) says that the total consumption at date 1 in each state η must be less than or equal to the supply of the good (equals the amount of the short asset). Condition (5) is an inequality because it is possible to transform an excess of the short asset at date 1 into consumption at date 2. Condition (6) says that the sum of consumption over the two periods is equal to the total returns from the two assets. Alternatively, we can read this as saying that consumption at date 2 is equal to the return on the long asset *plus* whatever is left over from date 1.

A *pure intermediated equilibrium* consists of a price system (p, q) and an attainable allocation $\{(x_i, y_i)\}$ such that, for every intermediary i, the choice of mechanism x_i and portfolio y_i solves the decision problem

$$\begin{array}{ll} \max & \sum_\eta \sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i(\theta_i, \eta), \theta_i) \\ \text{s.t.} & x_i \in X_i; \\ & \sum_\eta q(\eta) \sum_{\theta_i} \lambda(\theta_i, \eta) p(\eta) \cdot x_i(\theta_i, \eta) \leq \sum_\eta q(\eta) p(\eta) \cdot (y_i, (\mu_i - y_i)R(\eta)). \end{array} \tag{7}$$

In this specification, we assume that the definition of incentive compatibility is given by X_i. Competition and free entry force intermediaries to maximize the welfare of the typical depositor.

In a pure equilibrium, we assume that all intermediaries serving type i choose the same portfolio and contract. To ensure the existence of equilibrium, we need to allow for the possibility that intermediaries of type i make different choices. A *mixed allocation* is defined by a finite set of numbers $\{\rho^j\}$ and allocations $\{(x_i^j, y_i^j)\}$ such that $\rho^j \geq 0$ and $\sum_j \rho^j = 1$. A mixed allocation is *attainable* if the mean $\{(x_i, y_i)\} = \sum_j \rho^j \left\{\left(x_i^j, y_i^j\right)\right\}$ satisfies the market-clearing conditions (5) and (6). Note that $\{(x_i, y_i)\}$ may not be an allocation: even if each $\left(x_i^j, y_i^j\right)$ belongs to X_i the mean (x_i, y_i) may not because the set X_i is not convex.

A *mixed intermediated equilibrium* consists of a price system (p, q) and a mixed attainable allocation $\left\{\left(\rho^j, x_i^j, y_i^j\right)\right\}$ such that for every intermediary i, and every subgroup j, the choice $\left(x_i^j, y_i^j\right)$ solves the decision problem (7).

Assumption 1: (Non-empty interior) For any ex ante type $i = 1, ..., n$, for any consumption bundle $x_i \in X_i$ and price system $p \in P$, and for any $\varepsilon > 0$, either

$$\sum_\eta p(\eta) \cdot \left(\sum_{\theta_i} \lambda(\theta_i, \eta) x_i(\theta_i, \eta) \right) = 0$$

or there exists a bundle x_i' within a distance ε of x_i such that

$$\sum_\eta p(\eta) \cdot \left(\sum_{\theta_i} \lambda(\theta_i, \eta) x_i'(\theta_i, \eta) \right) < \sum_\eta p(\eta) \cdot \left(\sum_{\theta_i} \lambda(\theta_i, \eta) x_i(\theta_i, \eta) \right).$$

Theorem 1 *Under the maintained assumptions, if Assumption 1 is satisfied, there exists a mixed intermediated equilibrium.*

A pure equilibrium is a special case of a mixed equilibrium, but there is no guarantee that pure equilibria exist. In fact, we show in Section 7 that pure equilibria fail to exist in straightforward cases. Mixed equilibria are interesting in their own right. Among other things, they demonstrate the efficiency of having ex ante identical banks choose different risk strategies. They also illustrate another dimension of risk taking behavior. Two economies may appear similar in terms of the average level of risk in intermediaries' portfolio, but one may have every bank choosing a moderate level of risk while in the other most banks choose a very safe strategy while a few choose an extremely risky one. The behavior of the two economies may be very different.

For some purposes it is useful to interpret a mixed intermediated equilibrium as a pure intermediated equilibrium with a different set of ex ante types. Let (i, j) denote the new ex ante sub-type consisting of investors in the subgroup j of ex ante type i and let $\mu_{ij} = \rho^j \mu_i$ denote the measure of investors in the new ex ante type (i, j). Notice that we have to define a distinct economy for every mixed intermediated equilibrium. This is because the weights μ_{ij} depend on the endogenous variables ρ^j. In a mixed equilibrium, all agents of given ex ante type i receive the same expected utility. Thus, the expected utility of sub-type (i, j) is the same as the expected utility of type (i, j') for any given type i. Taking the weights $\{\mu_{ij}\}$ as given, we might be able to find other pure equilibria of the artificial economy, but they would not necessarily be mixed equilibria of the original economy, because they would not necessarily satisfy this equilibrium condition. For most purposes, this is not an issue, so without loss of generality we can restrict attention to pure equilibria.

3.4 Efficiency

Whichever definition of intermediated equilibrium we choose, that is, whether we choose X_i or X_i^* as the set of incentive-compatible contracts, we can show that the equilibrium allocation is incentive-efficient. An attainable allocation $(x, y) = \{(x_i, y_i)\}$ is *incentive efficient* if there does not exist an attainable mixed allocation $\left\{\left(\rho^j, x_i^j, y_i^j\right)\right\}$ such that

$$\sum_{\eta}\sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i^j(\theta_i, \eta), \theta_i) \geq \sum_{\eta}\sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i(\theta_i, \eta), \theta_i)$$

for every (i, j) with strict inequality for some (i, j). (This definition differs from Pareto efficiency only to the extent that we restrict attention to the incentive-compatible mechanisms $x_i \in X_i$).

In order to prove the incentive-efficiency of equilibrium, we need an additional regularity condition:

Assumption 2: (Local non-satiation) For any ex ante type $i = 1, ..., n$, for any mechanism $x_i \in X_i$, and for any $\varepsilon > 0$, there exists a mechanism $x_i' \in X_i$ within a distance ε of x_i such that

$$\sum_{\eta}\sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i'(\theta_i, \eta), \theta_i) > \sum_{\eta}\sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i(\theta_i, \eta), \theta_i).$$

Remark: This assumption is the counterpart of the non-satiability assumptions used in the classical theorems of welfare economics. It requires more than the non-satiability of each ex post type's utility function, $u_i(\cdot, \theta_i)$, however because the incentive constraints must be satisfied also. To illustrate the meaning of Assumption 2, consider the following example. There are two ex post types $\theta_i = 1, 2$ with utility functions $u_i(c, 1)$ and $u_i(c, 2)$. The utility function are defined as follows:

$$u_i(c, 1) = c_1 + c_2,$$

$$u_i(c, 2) = c_1 + c_2 \text{ if } c_1 + c_2 \leq 1$$

and for any utility level $\bar{u} > 1$, we denote the indifference curve consisting of the locus of consumption bundles yielding $\bar{u}$ to an agent of type $\theta_i = 2$ by $I_i(\bar{u}, 2)$ and define it by putting

$$I_i(\bar{u}, 2) = \left\{ (c_1, c_2) = \left((1 - \alpha) \frac{\bar{u} + 1}{2}, \alpha \bar{u} \right) : 0 \leq \alpha \leq 1 \right\}.$$

Notice that both types have linear indifference curves, but their indifference curves have different slopes at consumption bundles $c = (c_1, c_2)$ such that $c_1 + c_2 > 1$. Now consider an incentive-compatible mechanism x_i satisfying $x_i(1, \eta) = (1, 0)$ and $x_i(2, \eta) = (0, 1)$. Both incentive constraints (1) are just satisfied. Any mechanism x_i' that is ε-close to x_i and makes type i better off ex ante must make at least one of the ex post types better off and the incentive constraint then requires that the other ex post type be better off too. In other words, both of the consumption bundles $x_i'(1, \eta)$ and $x_i'(2, \eta)$ must lie above the line $c_1 + c_2 = 1$. Above the line $c_1 + c_2 = 1$, the two ex post types have linear indifference curves and the indifference curve of type 2 is steeper than that of ex post type 1. Since $x_i'(\theta_i, \eta)$ is very close to $x_i(\theta_i, \eta)$ for $\theta_i = 1, 2$ at least one type θ_i must envy the other. Thus, local non-satiation is not satisfied.

Theorem 2 *Under the maintained assumptions, if (p, q, x, y) is an intermediated equilibrium and Assumption 2 is satisfied, then the allocation (x, y) is incentive-efficient.*

Proof. See Section **??**. ■

Remark: The definition of incentive efficiency used here assumes that each investor of type i has been assigned to group (i, j) when we make the welfare comparison between the two allocations. A weaker requirement, leading to a stronger notion of incentive efficiency, assumes that investors of type i are randomly assigned to subgroups (i, j) and identifies their welfare with their expected utility before assignment. In that case, the payoffs in the inequality above would be multiplied by ρ^j and summed over subgroups (i, j) for each i. Jensen's inequality implies that the theorem remains true with this definition. However, if we want to take this approach, then we should really define incentive efficiency in terms of the true ex ante types, not the ex ante types of the artificial economy.

Theorem 2 is in the spirit of Prescott and Townsend (1984a, b), but the present model takes the decentralization of the incentive-efficient allocation a step further. Markets are used in Prescott and Townsend (1984a, b) to allocate mechanisms to agents at the first date. After the first date all trade is intermediated by the mechanism. Here, markets are also used for sharing risk and for intertemporal smoothing and intermediaries are active participants in markets at each date.

Remark: The incentive-efficiency of equilibrium is in marked contrast to the results in Bhattacharya and Gale (1986). The difference is explained by the informational assumptions. In the model above, there is no asymmetry of information in financial markets. Once the state η is observed, all aggregate uncertainty is resolved. The distribution of ex post types in each intermediary is a function of η and hence becomes common knowledge once η is revealed. Trading Arrow securities at date 0 is sufficient to provide optimal insurance against all aggregate shocks at date 1. In Bhattacharya and Gale (1986), by contrast, an intermediary's true demand for liquidity is private information at date 1. Markets for Arrow securities cannot provide incentive-efficient insurance against private shocks.

While symmetry of information in financial markets is a useful benchmark, one can easily imagine circumstances in which intermediaries have private information, for example, the intermediary knows the distribution of ex post types among its depositors, but outsiders do not. In that case, providing incentive-efficient insurance to the intermediaries would require us to supplement markets for Arrow securities with an incentive-compatible insurance mechanism, as in Bhattacharya and Gale (1986).

4 Optimal banking

In the benchmark model defined in Section 3, intermediaries use general, incentive-compatible contracts. In reality, we do not observe such complex contracts. In this section, we assume that financial institutions take the form of *banks*, which are required to use demand deposit contracts.

The use of demand deposits reduces welfare for two reasons. First, it distorts the intermediaries' choice of the investment portfolio and the risk sharing contract. Second, it leads to the possibility of bank runs, which cause further distortion of risk sharing through premature liquidation of bank assets. However, there is no market failure. Our model treats the use of deposit contracts as a constraint on banks, the result of transaction costs, legal restrictions, asymmetric information or problems of verifiability. As long as the government is subject to similar restrictions, it cannot do better. In fact, we can show that, under the maintained assumptions, the allocation of risk sharing is *constrained-efficient*. This is a very strong result. The use of demand deposits may reduce welfare compared to the benchmark model with unrestricted (incentive-compatible) contracts, but it does not provide a rationale for government regulation or intervention.

In the original Diamond-Dybvig model, there is no aggregate (intrinsic) uncertainty and a deposit contract is interpreted to be a promise to pay a fixed amount d_t at date t. More generally, a deposit contract offers investors a fixed one-period rate of return on the amount of the good left on deposit. When there is aggregate uncertainty, the interest rate on deposits changes over time in response to new information. As a result, depositors face uncertainty from two sources. First, the rate of return promised at date 1 is contingent on the aggregate state η and appears stochastic from the point of view of date 0. Second, the bank may be unable to make the payment promised at date 0 either because the returns to the long asset are too low or the liquidity demands of the depositors are too great.

An investor of type i deposits his endowment of one unit of the good with the bank at date 0. In exchange, he is promised a fixed return of $\bar{r}_{i0}$ at date 1. He can withdraw any amount between 0 and $\bar{r}_{i0}$ at date 1. The one-period interest rate is reset at date 1 after the true state η is observed. Whatever is left in the account between date 1 and date 2 will earn a fixed return of $r_{i1}(\eta)$. In the event that the bank cannot make the promised payment $\bar{r}_{i0}$ at date 1, the value of the deposit at date 1 is denoted by $r_{i0}(\eta) \leq \bar{r}_{i0}$.

If the bank (partially) defaults on its promised payments, then it must liquidate all of its assets in order to make the maximum possible payment to the investors at date 1. It follows that all the investors will withdraw the maximum amount at date 1 and use the short asset to allocate their consumption optimally over the remaining dates $t = 1, 2$. Note that the value of the bank's assets is determined by the market prices, independently of any liquidation decision. By contrast, in the model of Diamond and Dybvig (1983), goods are actually used up in the process of liquidating the long asset. Here, the only cost, ex post, arises from the fact a bank default excludes depositors from using the asset market to provide consumption at date 2. This may result in a distortion of intertemporal marginal rates of substitution.

Note also that unlike the case of general intermediaries, where it was largely a matter of indifference whether we allowed investors to hold the short asset, here it is essential to allow investors to do so. If a bank closes down at date 1, the only way for investors to provide themselves with consumption at date 2 is to hold the short asset.

The foregoing considerations are summed up in the following definition. A *deposit contract* is defined by a contingent payment schedule $w_i : H \to \mathbf{R}_+$ and a price schedule $\pi_i : H \to [0, 1]$ satisfying the constraint

$$\left[w_i(\eta) < \bar{w}_i \equiv \max_{\eta} w_i(\eta) \right] \Longrightarrow [\pi_i(\eta) = 1] \,. \tag{8}$$

At date 1 in state η, the investor is offered a choice from the budget set

$$\{c \in \mathbf{R}^2_+ : (1, \pi_i(\eta)) \cdot c \leq w_i(\eta)\}$$

where $c = (c_1, c_2)$ is the consumption vector. The restriction $\pi_i(\eta) \leq 1$ reflects the fact that investors have access to the short asset. If the return promised by the bank were less than the return on the short asset, investors would simply take the entire amount $w_i(\eta)$ out of the bank and invest it in the short asset. Thus, there is no loss of generality in assuming that, in the previous notation, $r_{i1}(\eta) \geq 1$ or, in the current notation, $\pi_i(\eta) \leq 1$. In the event that the bank cannot pay the maximum promised amount $\bar{w}_i \equiv \max_\eta w_i(\eta)$, the condition (8) says that all investors will want to withdraw the entire amount $w_i(\eta)$ from the bank and invest it in the short asset instead. Thus, there is no loss of generality in setting $\pi_i(\eta) = 1$ in those states.

Faced with a deposit contract (w_i, π_i), a depositor of type i will choose a consumption bundle x_i such that

$$x_i(\theta_i, \eta) \in \arg\max\{u_i(c, \theta_i) | (1, \pi_i(\eta) \cdot c \leq w_i(\eta)\}$$

for each type θ_i and each state η. Let $\hat{X}_i$ denote the set of mechanisms that can be generated by an appropriate deposit contract. It is clear that every mechanism $x_i \in \hat{X}_i$ is incentive compatible, so $\hat{X}_i \subseteq X_i$.

A *pure banking allocation* $(x, y) = \{(x_i, y_i)\}$ consists of a mechanism $x_i \in \hat{X}_i$ and a portfolio $0 \le y_i \le \mu_i$ for each type i. The banking allocation (x, y) is *attainable* if it satisfies the market-clearing conditions (5) and (6). A *pure banking equilibrium* consists of a price system (p, q) and an attainable banking allocation (x, y) such that, for each type of investor i, (x_i, y_i) maximizes the depositor's expected utility:

$$\begin{array}{ll} \max & \sum_\eta \sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i(\theta_i, \eta), \theta_i) \\ \text{s.t.} & x_i \in \hat{X}_i; \\ & \quad \sum_\eta q(\eta) \sum_{\theta_i} \lambda(\theta_i, \eta) p(\eta) \cdot x_i(\theta_i, \eta) \\ & \quad \le \sum_\eta q(\eta) p(\eta) \cdot (y_i, (\mu_i - y_i) R(\eta)). \end{array} \tag{9}$$

$>$From inspection, this decision problem is identical to the intermediary's except for one thing: the set of feasible mechanisms X_i is replaced by the set $\hat{X}_i \subseteq X_i$ consisting of mechanisms that can be implemented through deposit contracts.

A *mixed banking allocation* is defined by a finite set of numbers $\{\rho^j\}$ and banking allocations $\{(x_i^j, y_i^j)\}$ such that $\rho_j \ge 0$ and $\sum_j \rho^j = 1$. A mixed banking allocation is *attainable* if the mean $\{(x_i, y_i)\} = \sum_j \rho^j \left\{\left(x_i^j, y_i^j\right)\right\}$ satisfies the market-clearing conditions (5) and (6).

A *mixed banking equilibrium* consists of a price system (p, q) and a mixed attainable allocation $\left\{\left(\rho^j, x_i^j, y_i^j\right)\right\}$ such that for every intermediary i, and every subgroup j, the choice $\left(x_i^j, y_i^j\right)$ solves the decision problem (9). To ensure existence of a mixed banking equilibrium we need an interiority assumption.

ASSUMPTION 1A: (Non-empty interior) For any ex ante type $i = 1, ..., n$, for any consumption bundle $x_i \in \hat{X}_i$ and price system $p \in P$, and for any $\varepsilon > 0$, either

$$\sum_\eta p(\eta) \cdot \left(\sum_{\theta_i} \lambda(\theta_i, \eta) x_i(\theta_i, \eta) \right) = 0$$

or there exists a bundle x_i' within a distance ε of x_i such that

$$\sum_\eta p(\eta) \cdot \left(\sum_{\theta_i} \lambda(\theta_i, \eta) x_i'(\theta_i, \eta) \right) < \sum_\eta p(\eta) \cdot \left(\sum_{\theta_i} \lambda(\theta_i, \eta) x_i(\theta_i, \eta) \right).$$

Note that even though $\hat{X}_i \subset X_i$, Assumption 1A is not implied by Assumption 1, because $x'_i \in X_i$ does not imply $x'_i \in \hat{X}_i$.

Theorem 3 *Under the maintained assumptions, there exists a mixed banking equilibrium if Assumption 1A is satisfied.*

As in the case of the intermediated equilibrium, we can re-interpret mixed banking equilibria as (pure) banking equilibria for an artificial economy with more ex ante types. So, without loss of generality we restrict attention to pure equilibria in what follows.

Because they are restricted to deposit contracts, it is clear that banks cannot always do as well as general intermediaries. However, the banking equilibrium is constrained-efficient in the sense that no feasible banking allocation can make every type of depositor better off. An attainable banking allocation (x, y) is *constrained-efficient* if there does not exist an attainable mixed banking allocation $\left\{\left(\rho^j, x_i^j, y_i^j\right)\right\}$ such that

$$\sum_\eta \sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i^j(\theta_i, \eta), \theta_i) \geq \sum_\eta \sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i(\theta_i, \eta), \theta_i)$$

for every (i, j) with strict inequality for some (i, j).

As in the case of the intermediated equilibrium, we need to assume local non-satiability in order to ensure efficiency.

ASSUMPTION 2A: (Local non-satiation) For any ex ante type $i = 1, ..., n$, for any $x_i \in \hat{X}_i$, and for any $\varepsilon > 0$, there exists a mechanism $x'_i \in \hat{X}_i$ within a distance ε of x_i such that

$$\sum_\eta \sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x'_i(\theta_i, \eta), \theta_i) > \sum_\eta \sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i(\theta_i, \eta), \theta_i).$$

Theorem 4 *If (p, q, x, y) is a banking equilibrium and Assumption 2A is satisfied, then the banking allocation (x, y) is constrained-efficient.*

Proof. The proof is essentially the same as for Theorem 2. ■

Remark: An important qualification to the constrained-efficiency of the banking equilibrium is the "equilibrium selection" implicit in the definition of equilibrium. Banks are price takers so equilibrium prices and hence investors' behavior at dates 1 and 2 are perceived to be independent of the choices made by banks at date 0. In fact, once we recognize the possibility of multiple equilibria and the dependence of equilibrium selection at date 1 in a possibly complicated way on the banks' choices at date 0, it is clear that there are many equilibria, some of which are inefficient. This aspect of the problem is beyond the scope of this paper, but it deserves to be investigated in the future.

5 Complete participation and redundancy

Cone (1983) and Jacklin (1986) have pointed out that the beneficial effects of banks in the Diamond-Dybvig (1983) model depend on the assumption that there are no markets in which assets can be traded at the intermediate date. If such markets exist, and depositors are allowed to participate in the markets, then the market allocation weakly dominates the allocation implemented by the banks. Banks are redundant in the sense that they cannot improve on the risk sharing achieved by markets alone. A similar result holds here. The ability of intermediaries to provide insurance against shocks to liquidity preference depends crucially on the assumption that investors cannot participate directly in asset markets. Increasing access to financial markets actually lowers welfare. To show this, we first characterize an equilibrium with markets but without intermediaries. Then we show that the introduction of intermediaries is redundant.

The market data is the same as in Section 2. Investors are allowed to trade in markets for Arrow securities at date 0 and can trade in the spot markets for the good at date 1. Let y_i denote the portfolio chosen by a representative investor of type i and let $x_i(\theta_i, \eta)$ denote the consumption bundle chosen by an investor of type θ_i in state η. The investor's choice of $(x_{i,}y_i)$ must solve the decision problem:

$$\begin{array}{ll} \max & \sum_\eta \sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i(\theta_i, \eta), \theta_i) \\ \text{s.t.} & \sum_\eta q(\eta) \cdot x_i(\theta_i, \eta) \leq \sum_\eta q(\eta) p(\eta) \cdot (y_i, (1 - y_i) R(\eta)), \forall \theta_i. \end{array}$$

The maximization problem for the individual is similar to that of an intermediary of type i except for two things. First, there is no explicit incentive constraint $x_i \in X_i$. Second, there is no insurance provided against the realization of θ_i. This is reflected in the fact that, instead of summing the budget constraint over η and θ_i, it is summed over η only and must be satisfied for each θ_i. The ex ante type i can redistribute wealth across states η in any way he likes, but each type θ_i will get the same amount to spend in state η. As a result, for each realization of θ_i in a given state η, the choice of $x_i(\theta_i, \eta)$ must maximize the utility function $u_i(x_i(\theta_i, \eta), \theta_i)$ subject to a budget constraint that depends on η but not on θ_i. This ensures that incentive compatibility is satisfied.

An equilibrium for this economy consists of a price system $\{p, q\}$ and an attainable allocation $\{(x_i, y_i)\}$ such, that for every type i, (x_i, y_i) solves the individual investor's problem above.

Comparing this definition of equilibrium with the intermediary equilibrium, it is clear that an equilibrium with complete market participation cannot implement the incentive-efficient equilibrium except in the special case where there are no gains from sharing risk against liquidity preference shocks.

The more interesting question is whether introducing intermediaries in this context can make investors any better off. The answer, as we have suggested, is negative. To see this, note first of all that in an equilibrium with complete participation, all that investors care about in each state is the market value of the bundle they receive from the intermediary. If an intermediary offers a mechanism x_i the investors of type θ_i will report the ex post type $\hat{\theta}_i$ that maximizes $p(\eta) \cdot x_i(\hat{\theta}_i, \eta)$ in state η. Using the Revelation Principle, there is no loss of generality in assuming that the value of consumption $p(\eta) \cdot x_i(\theta_i, \eta) = w_i(\eta)$ is independent of θ_i. But this means that the intermediary can do nothing more than replicate the effect of the markets for Arrow securities. More precisely, the intermediary is offering a security that pays $w_i(\eta)$ in state η and the intermediary's budget constraint

$$\sum_{\eta} q(\eta) w_i(\eta) \leq \sum_{\eta} q(\eta) p(\eta) \cdot (y_i, (\mu_i - y_i) R(\eta))$$

ensures that $\{w_i(\eta)\}$ can be reproduced as a portfolio of Arrow securities at the same cost.

Theorem 5 *In an equilibrium with complete market participation, the introduction of intermediaries is redundant in the sense that an equilibrium allocation in an economy with intermediaries is weakly Pareto-dominated by an equilibrium allocation of the corresponding economy without intermediaries.*

This is, essentially, a "Modigliani-Miller" result, saying that whatever an intermediary can do can be done (and undone) by the market.

The same result holds a fortiori if we replace the general intermediary with a bank issuing deposit contracts. As before, complete participation implies that investors only care about the market value of the bundle they receive from the bank. In an economy with market participation and no banks, every investor of type i receives the same amount in state η, independently of his ex post type θ_i. Likewise, in an economy with banks, every depositor of type i receives the value of the bank's portfolio in state η, independently of his ex post type θ_i. It is a matter of indifference whether the bank formally defaults or not (the depositor receives the same amount in either case), so the value of the fixed payment d is irrelevant and might as well be set equal to zero.

6 Prudential regulation

The absence of markets for insuring individual liquidity shocks does not by itself lead to market failure. As we showed in Section 3, intermediaries can achieve an *incentive-efficient* allocation of risk under certain conditions. Under the same conditions, banks can achieve a *constrained-efficient* allocation of risk. Two assumptions are crucial for these results: markets for aggregate uncertainty are complete (there is an Arrow security for each state η) and market participation is limited. Markets for aggregate risk allow intermediaries to share risk among different ex ante types of investor and limited participation allows intermediaries to offer insurance against the realization of ex post types.

In order to generate a market failure and provide some role for government intervention, an additional friction must be introduced. Here we assume that it comes in the form of *incomplete markets for aggregate risk.* There are two sources of aggregate risk in this model. One is the return to the risky asset $R(\eta)$; the other is the distribution of investors' liquidity preferences $\lambda(\theta_i, \eta)$. Although there is a continuum of investors, correlated liquidity shocks give rise to aggregate fluctuations in the demand for liquidity.

Market incompleteness may be expected to vary according to the source of the risk. On the one hand, asset returns can be hedged using markets for stock and interest rate options and futures. For example, financial markets allow intermediaries to take hedging positions on future asset prices and interest rates. Additional hedges can be synthesized using dynamic trading strategies. On the other hand, liquidity shocks seem harder to hedge. Asset prices and interest rates are functions of the *aggregate* distribution of liquidity shocks in the economy. Inverting this relationship, we can use asset prices and interest rates to infer the average demand for liquidity, but not the liquidity shock experienced by a particular intermediary. Hence, trading options and futures will not allow intermediaries to provide insurance against their own liquidity shocks. So markets for hedging liquidity shocks are likely to be more incomplete than the markets for hedging asset returns.

In the sequel, we analyze a polar case: we assume that there exist complete markets for hedging asset return shocks, but no markets for hedging liquidity shocks. Since complete markets for asset-return shocks permit efficient sharing of that risk, we might as well assume that asset returns are non-stochastic. In what follows, we adopt the simplifying assumption that the return to the long asset is a constant $r > 1$ and the only aggregate uncertainty comes from liquidity shocks.

6.1 Regulation of intermediaries

We continue to assume that each intermediary serves a single type of investor. Let x_i denote the mechanism and y_i the portfolio chosen by the representative intermediary of type i. The problem faced by the intermediary is to choose (x_i, y_i) to maximize the expected utility of the representative depositor subject to multiple, state-contingent budget constraints:

$$\begin{array}{ll} \max & \sum_{\eta} \sum_{\theta_i} \lambda(\theta_i, \eta) u_i(x_i(\theta_i, \eta), \theta_i) \\ \text{s.t.} & x_i \in X_i; \\ & \sum_{\theta_i} \lambda(\theta_i, \eta) p(\eta) \cdot x_i(\theta_i, \eta) \\ & \leq p(\eta) \cdot (y_i, (\mu_i - y_i) R(\eta)), \forall \eta. \end{array}$$

The incompleteness of markets is reflected in the fact that there is a separate budget constraint for each aggregate state of nature η, rather than a single budget constraint that integrates surpluses and deficits across all states.

The work of Geanakoplos and Polemarchakis (1986) (hereafter GP) suggests that general equilibrium with incomplete markets (GEI) is generically inefficient. More precisely, GP show that a planner can make all investors better off by manipulating date 0 variables while relying on the market at date 1. Marginal rates of substitution vary across investor types when markets are incomplete; a change in the first-period allocation will change second-period prices in a way that causes transfers among investors in each state of nature at date 2; because of the difference in marginal valuations, state-contingent transfers can increase every agents' utility. The present model is not a special case of the GEI model, so the GP result does not apply. Still, the reasoning suggests that equilibrium will be inefficient in the sense of GP.

Rather than pursuing general inefficiency results, we investigate specific policy interventions to identify the precise ways in which welfare can be increased. Manipulation of portfolio choices at the first date can be interpreted as a form of *prudential regulation*. For example, a lower bound on the amount of the short asset held by intermediaries can be interpreted as a reserve requirement. We examine prudential regulation in the context of an economy with Diamond-Dybvig preferences.

As usual, there are n ex ante types of investors $i = 1, ..., n$. The ex ante types are symmetric and the size of each type is normalized to $\mu_i = 1$. There are two ex post types, called early consumers and late consumers. Early consumers only value consumption at date 1, while late consumers only value consumption at date 2. The ex ante utility function of type i is defined by

$$u_i(c_1, c_2, \theta_i) = (1 - \theta_i)U(c_1) + \theta_i U(c_2),$$

where $\theta_i \in \{0, 1\}$. The period utility function $U : \mathbf{R}_+ \to \mathbf{R}$ is twice continuously differentiable and satisfies the usual neoclassical properties, $U'(c) > 0, U''(c) < 0$, and $\lim_{c\searrow 0} U'(c) = \infty$.

Each intermediary is characterized at date 1 by the proportions of early and late consumers among its customers. Call these proportions the intermediary's type. Then aggregate risk is characterized by the distribution of types of intermediary at date 1. We assume that the cross-sectional distribution of types is constant and no variation in the total demand for liquidity. In this sense there is no aggregate uncertainty. However, individual intermediaries receive different liquidity shocks: some intermediaries will have high demand for liquidity and some will have low demand for liquidity. The asset market is used to re-allocate liquidity from intermediaries with a surplus of liquidity to intermediaries with a deficit of liquidity. It is only the cross-sectional distribution of liquidity shocks that does not vary. To make this precise, we need additional notation. We identify the state η with the n-tuple $(\eta_1, ..., \eta_n)$, where η_i is the fraction of investors of type i who are early consumers in state η. The fraction of early consumers takes a finite number of values, denoted by $0 < \sigma_k < 1$, where $k = 1, ..., K$. We assume that the marginal distribution of η_i is the same for every type i and that the cross-sectional distribution is the same for every state η. Formally, let λ_k denote the ex ante probability that the proportion of early consumers is σ_k. Let $H_{ik} = \{\eta \in H : \eta_i = \sigma_k\}$. Then $\lambda_k = \sum_{\eta \in H_{ik}} \lambda_i(0, \eta)$, for every k and every type i. Ex post, λ_k equals the fraction of ex ante types with a proportion σ_k of early consumers. Then $\lambda_k = n^{-1}\#\{i : \eta_i = \sigma_k\}$, for every k and every state η.

The return on the long asset is non-stochastic,

$$R(\eta) = r > 1, \forall \eta.$$

The only aggregate uncertainty in the model relates to the distribution of liquidity shocks across the different ex ante types i.

In what follows, we focus on symmetric equilibrium, in which the prices at dates 1 and 2 are independent of the realized state η and the choices of the intermediaries at date 0 are independent of the ex ante type i. Let the good at date 1 be the numeraire and let p denote the price of the good at date 2 in terms of the good at date 1. At date 0 the representative intermediary of type i chooses a portfolio $y_i = y$ and a consumption mechanism $x_i(\eta) = x(\eta_i)$ to solve the following problem:

$$\begin{array}{ll} \max & \sum_k \lambda_k \{\sigma_k U(x_1(\sigma_k)) + (1-\sigma_k) U(x_2(\sigma_k))\} \\ \text{s.t.} & \sigma_k x_1(\sigma_k) + (1-\sigma_k) p x_2(\sigma_k) \le y + pr(1-y), \forall k, \\ & x_1(\sigma_k) \le x_2(\sigma_k), \forall k. \end{array} \tag{10}$$

The incentive constraint $x_1(\sigma_k) \le x_2(\sigma_k)$ is never binding. To see this, drop the incentive constraint and solve the relaxed problem. The special early-consumer, late-consumer structure of preferences and the assumption $r > 1$ imply that the incentive constraint is automatically satisfied by the solution to the relaxed problem.

A symmetric equilibrium consists of an array (p, x, y) such that (x, y) solves the problem (10) for the equilibrium price p and the market-clearing conditions

$$\begin{array}{l} \sum_{k=1}^K \lambda_k \sigma_k x_1(\sigma_k) \le y, \\ \sum_{k=1}^K \lambda_k \{\sigma_k x_1(\sigma_k) + (1-\sigma_k) x_2(\sigma_k)\} = y + r(1-y). \end{array}$$

are satisfied.

Proposition 6 *There exists a unique symmetric equilibrium (p^*, x^*, y^*).*

Proof. *Uniqueness.* The Inada conditions imply that the optimal mechanism will provide positive consumption at both dates, so both assets are held in equilibrium. Both assets will be held between date 0 and date 1 only if the returns on the two assets are equal. If $pr > 1$ then the return on the long asset is greater than the return on the short asset between dates 0 and 1 and no one will hold the latter ; if $pr < 1$ then the return on the short asset is greater than the return on the long asset between dates 0 and 1 and no one will hold the long asset. Then equilibrium requires that $p = 1/r$ at date 1 (Allen and Gale (1994)).

At the price $p = 1/r$, the short asset is dominated between dates 1 and 2 so no one will hold it. It follows that all consumption at date 1 is provided by the short asset and all consumption at date 2 is provided by the long asset. Then the market-clearing conditions are

$$\begin{array}{l} \sum_{k=1}^{K} \lambda_k \sigma_k x_1(\sigma_k) = ny, \\ \sum_{k=1}^{K} \lambda_k (1-\sigma_k) x_2(\sigma_k) = nr(1-y). \end{array} \qquad (11)$$

At the price $p = 1/r$, intermediaries are indifferent between the two assets at date 0. Thus, the quantities of the assets held in equilibrium are determined by the market-clearing conditions.

Finally, note that at the price $p = 1/r$, the right hand side of the budget constraint in (10) is equal to 1 for all k, so the choice of x is independent of y. Strict concavity implies that there is at most one solution to the maximization problem. Hence, the equilibrium values are uniquely determined.

Existence. The existence of a symmetric equilibrium follows directly from the existence of a solution of the intermediary's decision problem (10). A solution exists because the choice set is compact and the objective function is continuous. The objective function is increasing, so the budget constraint holds as an equation for each k. Summing the budget constraint over k we get

$$\sum_k \lambda_k \left\{ \sigma_k x_1(\sigma_k) + (1-\sigma_k) p x_2(\sigma_k) \right\} = n,$$

where $p = 1/r$. Clearly, there exists a value of $0 \leq y \leq 1$ such that the market-clearing conditions (11) are satisfied. Then (p, x, y) constitutes a symmetric equilibrium. ■

To study the effect of regulation of intermediaries, we take as given the choice of portfolio y at date 0 and consider the determination of the consumption mechanism x and the market-clearing price p. Call (p, x, y) a regulated equilibrium if x solves the problem (10) for the given values of p and y, and (x, y) satisfies the market-clearing conditions (11).

Proposition 7 *For any p sufficiently close to p^*, there exists a unique regulated equilibrium (p, x, y).*

Proof. See Section **??**. ■

Suppose that (p, x, y) is a regulated equilibrium and $p > p^*$. Then the short asset is dominated by the long asset between dates 0 and 1 and no one will want to hold the short asset. If the regulator requires intermediaries to hold a minimum amount of the short asset then each intermediaries will want to hold the minimum. In that case, (x, y) solves the maximization problem (10) subject to the additional constraint $y \geq \bar{y}$ for an appropriately chosen value of $\bar{y}$. Similarly, if $p < p^*$ then the long asset is dominated and (x, y) solves the maximization problem subject to an additional constraint $y \leq \bar{y}$. Thus, any regulated equilibrium can be interpreted as an equilibrium in which the intermediaries choose the mechanism x and the portfolio y to maximize the expected utility of the depositors subject to a constraint that requires them to hold a minimum amount of the short asset in the case $p > p^*$ or of the long asset in the case $p < p^*$. In this sense, the regulator can implement the regulated equilibrium (p, x, y) by imposing an appropriate constraint $\bar{y}$ on the intermediaries' portfolio choice.

By manipulating the intermediaries' portfolio choice, the regulator is able to manipulate the equilibrium price p. The welfare effect of this change in price depends on the degree of risk aversion. We distinguish two cases, according to whether the degree of risk aversion is greater than or less than one. Let $W(p)$ denote the expected utility of the typical depositor in the regulated equilibrium (p, x, y).

Theorem 8 *For any $p > p^*$ sufficiently close to p^*, $W(p) > W(p^*)$ if*

$$-\frac{U''(c)c}{U'(c)} > 1. \tag{12}$$

For any $p < p^*$ sufficiently close to p^*, $W(p) > W(p^*)$ if

$$-\frac{U''(c)c}{U'(c)} < 1.$$

Proof. See Section **??**. ■

So an increase in liquidity (increase in the price of the long asset) may increase or decrease welfare, depending on the degree of risk aversion.

The intuition behind these results is the following. When relative risk aversion is high, optimal risk sharing requires intertemporal smoothing (the consumption at date 1 is higher than the present value of consumption at date 2). Then an increase in σ_k implies that per capita consumption must fall at both dates. However, the budget constraint ensures that total consumption must rise with σ_k at the first date. So per capita consumption and demand for liquidity go in opposite directions. An increase in liquidity (increase in the price of the long asset), reduces consumption in good states (where per capita consumption is high) and raises it in bad states (where per capita consumption is low). This transfer from high-consumption states to low-consumption states raises welfare. When relative risk aversion is low, the argument works in the opposite direction. The welfare effect depends on the correlation between equilibrium consumption and the liquidity shocks.

6.2 Bank regulation

A similar analysis can be carried out for an economy with banks. The analysis requires a different technique but the results turn out to be very similar. If the degree of relative risk aversion is less than one there is no possibility of bankruptcy in a symmetric banking equilibrium and welfare is improved by decreasing liquidity, i.e., forcing the banks to hold more of the *long* asset. Conversely, if the degree of relative risk aversion is greater than one, bankruptcy is possible in a symmetric banking equilibrium and welfare is improved by increasing liquidity, i.e., forcing the banks to hold more of the *short* asset. The details of the analysis can be found in Allen and Gale (2000c).

7 Examples

To illustrate the properties of the model, we examine a series of numerical examples in this section. First, we look at examples in which there is only uncertainty about the return to the long asset. Then we look at an example in which there is only uncertainty about liquidity shocks. In each case, we find a rich variety of equilibrium behavior.

7.1 Return shocks

We begin by considering an economy in which there is one ex ante type of investor. The decision problem of a representative intermediary is the same as the planner's problem, so the efficiency properties of equilibrium are immediate. Whether the incentive constraint is binding or not depends on the degree of uncertainty about the return to the long asset.

Example 1: We use the generalized Diamond-Dybvig preferences introduced in Section 6. There are two equally likely states $\eta = 1, 2$. In each state, an agent has an equal probability of being an early or a late consumer. The period utility function is $U(c) = Ln(c)$. The return on the long asset is contingent on the state:

$$R(1) = r; R(2) = 2.$$

By varying r, the return in state 1, we generate a variety of different equilibria.

General Intermediaries

The intermediaries receive a total of 1 unit of the good and invest y in the short term asset and $1-y$ in the long term asset. Intermediaries write general, incentive-compatible contracts with investors. Since there is only one ex ante type, the intermediary solves the same problem as the planner. Because of the log utility function, the first-best allocation takes a particularly simple form. The investment in the short asset is $y = 0.5$. Consumption at date 1 is independent of the state and the intermediary's holding of the short asset is just sufficient to provide the consumption at date 1. Consumption at date 2 is equal to returns to the long asset.

$$\begin{aligned} c_1(\eta) &= 1 = 2y; \\ c_2(\eta) &= 2R(\eta)(1 - y) = R(\eta). \end{aligned}$$

Example 1A. For high values of r the incentive constraint is automatically satisfied by the first-best allocation, that is, the first best is the same as the incentive-efficient allocation. This means that intermediaries can achieve the first best. More precisely, for $r \geq 1$, there exists a pure intermediated equilibrium that implements the first best. A typical equilibrium with $r = 1.5$ is shown in Table 1.

Example 1B. The incentive constraint is just satisfied by the first-best allocation for $r = 1$. For values of $r < 1$, the first-best allocation violates the incentive constraint so the incentive-efficient allocation is strictly worse than the first best in terms of welfare. For $0.4 \leq r < 1$, there is a unique pure intermediated equilibrium in which the consumption profile satisfies $c_1(1) < 2y$ and $c_2(2) = 2y$. In state 2, all of the short asset is consumed at date 1; in state 1, the returns to the long asset are so low that some of the short asset must be saved to provide consumption at date 2. Whenever a positive amount of the short asset is held between date 1 and date 2, efficiency requires that the intertemporal marginal rate of consumption be equal to 1. Thus, consumption is equalized across time in state 1. The amount invested in the short asset rises as the expected payoff on the long asset falls. Table 1 illustrates the equilibrium values for the case where $r = 0.5$.

Example 1C. It remains to consider the case $0 \leq r \leq 0.4$. As r continues to fall the intermediary invests more in the short asset and less in the long asset. The total output and consumption in state 2 are also falling and at $r = 0.4$ the consumption at date 1 is less than the holding of the short asset in both states: $c_1(\eta) < 2y$. Consumption is equalized across time in both states. The equilibrium values corresponding to $r = 0.3$ are given in Table 1.

In Example 1 there is only one ex ante type of investor and hence one type of intermediary. There is no possibility of risk sharing between intermediaries (types). The next example introduces a second ex ante type consisting of risk neutral investors. The purpose of this example is to illustrate how the Arrow securities allow cross-sectional risk sharing.

Example 2. There are two types of investors, the risk averse investors, denoted by A, and the risk neutral investors denoted by N. Otherwise things are the same as in Example 1. The period utility functions are $U_A(c) = Ln(c)$ and $U_N(c) = c$, respectively. The measure of each type is normalized to 1 and each type has a probability 0.5 of being an early or late consumer.

As in Example 1, the qualitative features of the equilibrium depend on whether some of the short asset is carried over from date 1 to date 2. One case is sufficient to illustrate how the Arrow securities allow risk to be shared. For $r \geq 0.4$ the allocation is the same as in Example 1A, except that all uncertainty is absorbed by the risk neutral type. The equilibrium portfolios are $y_A = 0.5$ and $y_N = 0$, for the representative intermediaries of type A and type N respectively.

Table 1 gives the equilibrium values for the case where $r = 0.5$. The market-clearing prices are $q(1) = 0.5, q(2) = 0.5, p_2(1) = 0.8, p_2(2) = 0.8$. The portfolio returns (receipts) of each intermediary and consumption of each type are given in the following table.

	State 1	State 2
$(y_A, R(\eta)(1 - y_A))$	$(1, 0.5)$	$(1, 2)$
$(y_N, R(\eta)(1 - y_N))$	$(0, 1)$	$(0, 4)$
$(c_{A1}(\eta), c_{A2}(\eta))$	$(1, 1.25)$	$(1, 1.25)$
$(c_{N1}(\eta), c_{N2}(\eta))$	$(0, 0.25)$	$(0, 4.75)$

Risk sharing between the two groups is achieved through trading in the Arrow-security markets at date 0. These allow the type A intermediaries to buy 0.75 per capita in state 2 at date 2. They balance their budget by selling 0.75 per capita in state 2 at date 2. The type N's are willing to take the other side of these trades since they are risk neutral. It can be seen that this improvement in risk sharing increases the expected utility of type A's from 0.05 in Example 1B to 0.11 here.

When there are two groups each type of intermediary has its own objective function. It is no longer immediate that that the allocation is incentive-efficient. However, the results of Section 3 demonstrate that the invisible hand of the market does indeed work here and the allocation is incentive-efficient.

There are other intermediated equilibria corresponding to these parameter values. In particular, the portfolio holdings of the type-A intermediaries are not uniquely determined. For example, the type-A intermediaries could hold less of the short term asset and more of the long term asset, as long as this is offset by changes in the holdings of the type-N intermediaries. As long as the aggregate portfolio remains the same, investors' consumption and welfare is unchanged.

Banks

Now we assume that the financial institutions are banks, that is, they are restricted to issuing deposit contracts. A deposit contract promises a fixed payment at date 0. If the bank is unable to meet its commitment to its depositors, then it is liquidated and all its asset are sold.

Example 3. To illustrate the properties of a banking equilibrium, we revert to the parameters from Example 1.

Example 3A. For $r \geq 1$ the first best can again be implemented as a banking equilibrium. The first-best allocation requires that consumption at date 1 be the same in both states, so the optimal incentive-compatible contract is in fact a deposit contract. Moreover, bankruptcy does not occur in equilibrium. The allocation corresponding to a pure banking equilibrium is exactly the same as the allocation corresponding to a pure intermediated equilibrium, which as we have seen is the same as the first best. For $r = 1.5$, for example, the equilibrium values will be the same as for Example 1A, as shown in Table 1.

The main difference between Examples 1A and 3A is that the prices supporting the first best allocation are unique in Example 1A but are not in Example 3A. In Example 1A the general intermediaries can freely vary payoffs across states and dates so only one set of prices can support the optimal allocation. Because the banks are restricted to using deposit contracts in Example 3A, consumption at date 1 is independent of the state as long as there is no bankruptcy. This prevents the banks from arbitraging between the states at date 1. In effect, we have lost one equilibrium as condition compared with the intermediated equilibrium. Any vector of prices $(q(1), q(2), p_2(1), p_2(2))$ that supports the equilibrium allocation at date 2 is consistent with a banking equilibrium. The following no-arbitrage conditions are satisfied

$$
\begin{aligned}
q(1)p_2(1) &= 0.33 && (13)\\
q(2)p_2(2) &= 0.25 && (14)\\
q(1) + q(2) &= 1 && (15)
\end{aligned}
$$

and the existence of the short asset implies that

$$p_2(1) \leq 1; p_2(2) \leq 1.$$

Combining these conditions, it follows that the range of possible prices for $q(1)$ is

$$0.33 \leq q(1) \leq 0.75.$$

The other prices are determined by (13)-(15).

For $r < 1$ deposit contracts cannot be used to implement the first best or incentive-efficient allocation. In particular, a deposit contract that is consistent with the incentive-efficient consumption in state 2 would imply bankruptcy in state 1 (see, e.g., Example 1B). If all banks were to offer a contract which allowed bankruptcy in state 1 they would all be forced to liquidate their assets. If they all do this, there is no bank on the other side of the market to buy the assets and the price of future consumption, $p_2(1)$, falls to 0. This cannot be an equilibrium because it would be worthwhile for an individual bank to offer a contract which did not involve bankruptcy. This bank could then buy up all the long term asset in state 1 at date 1 for $p_2(1) = 0$ and make a large profit.

This argument indicates that for $r < 1$ there cannot be a pure banking equilibrium with bankruptcy. There are two other possibilities. One is that there is a pure equilibrium in which all banks remain solvent. The other is that there is a mixed equilibrium in which banks choose different contracts and portfolios. In particular, a group of banks with measure ρ makes choices that allow it to remain solvent in state 1 while the remaining $1 - \rho$ banks go bankrupt. The banks that are ensured of solvency are denoted type S while those that can go bankrupt are denoted type B. It turns out that both the pure equilibrium with all banks remaining solvent and the mixed equilibrium can occur when $r < 1$.

Example 3B. For $0.90 \leq r < 1$ there is a pure banking equilibrium in which all the banks choose to remain solvent ($\rho = 1$). There is again a range of prices which can support the allocation. Type-S banks stay solvent by lowering the amount promised at date 1, lowering their investment in the short asset and increasing their investment in the long asset. The equilibrium values for the case where $r = 0.95$ are shown in Table 1. The range of $q(1)$ that will support the allocation and prevent entry by type B banks is $0.55 < q(1) < 0.77$. The other prices are determined by the equilibrium conditions

$$\begin{aligned} q(1)p_2(1) &= 0.54 \\ q(2)p_2(2) &= 0.24 \\ q(1) + q(2) &= 1. \end{aligned}$$

For $0.90 \leq r < 1$, there is no equilibrium with a positive measure of type-B banks. In this case the prices at date 0 are uniquely determined at $q(1) = 2/3$ and $q(2) = 1/3$ and at these prices the type B depositors are worse off than the type S depositors.

Example 3C. For $0 < r < 0.90$ there exists a mixed banking equilibrium. At the boundary $r = 0.90$, the set of prices that will support the equilibrium of the type shown in Example 3B shrinks to a singleton: $(q(1) = 2/3, q(2) = 1/3)$. For values of $r < 0.90$ but close to 0.90 the proportion of type-S banks is large ($\rho \approx 1$). As r falls, the proportion of type-B banks rises until as r goes to 0 the proportion of type-B banks $1 - \rho$ approaches 1. The mixed banking equilibrium corresponding to $r = 0.7$ is shown in Table 1. The proportion of type-S banks is $\rho = 0.63$. As in the earlier examples, the portfolios of individual banks are indeterminate. It is only the aggregate portfolio that is determined by the equilibrium conditions. Liquidity is transferred between states by the use of Arrow securities.

There are a number of other types of equilibrium, depending on whether the short asset is held between date 1 and date 2. The properties of these equilibria are similar to those illustrated in Examples 1 and 2 and will not be repeated here.

The mixed banking equilibrium illustrates a number of interesting phenomena. First, banks serving identical depositors may do quite different things in equilibrium. One subset will find it optimal to follow a safe strategy and avoid the risk of bankruptcy. Another subset will follow a risky strategy that makes bankruptcy inevitable in state 1. As a result, even when all investors are identical ex ante, there can be a financial crisis in which some banks remain solvent while others go under. In the context of the example, the crisis is more severe when the state-1 return r is lower. However, the results in Section 4 show that these crises are consistent with the constrained efficiency of equilibrium. There is no justification for intervention or regulation of the financial system. A planner could not do better.

Many analyses of banking assume that there is a technology for early liquidation of the long term asset. The resulting equilibria are typically symmetric. This example illustrates that this assumption materially changes the form of the equlibrium. With endogenous liquidation one group of banks must provide liquidity to the market and this results in mixed equilibria.

A comparison of Example 3C with Example 1 shows the importance of deposit contracts in causing financial crises when attention is restricted to essential crises. In Example 1 the ability to use state contingent contracts with investors means that bankruptcy never occurs and as a result there is never a crisis. However, with deposit contracts crises of varying severity always occur in the low output state.

7.2 Liquidity shocks

The crises in these examples are generated by shocks to asset returns. Similar results can be obtained by assuming aggregate uncertainty about the demand for liquidity.

Example 4. This example is similar to Example 1. There are two main differences. First, it is assumed that there is no uncertainty about the return to the long asset:

$$R(1) = R(2) = r.$$

Second, there is aggregate uncertainty about the demand for liquidity. The states $\eta = 1, 2$ are equally likely but now the proportions of early and late consumers differ across states:

Proportion	State 1	State 2
Early consumers	0.4	0.6
Late consumers	0.6	0.4

By varying r we can generate the same phenomena as in the case of return shocks. Rather than go through all the different cases, we illustrate a pure intermediated equilibrium for the case $r = 1.6$. The equilibrium values are shown in Table 1. For this case the aggregate liquidity constraint binds in both states because the return on the long asset is high. At date 1 all the proceeds from the short asset are used for consumption. In state 1 aggregate liquidity needs are low at date 1. Since consumption is split evenly among early consumers per capita consumption is high compared to state 2 where aggregate liquidity needs are high at date 1. At date 2 there are many late consumers in state 1 so consumption per capita is lower than in state 2 where there are relatively few late consumers.

8 Proofs

8.1 Proof of Proposition 2

The argument follows the lines of the familiar proof of the first fundamental theorem of welfare economics. Let (x, y, p, q) be an equilibrium and suppose, contrary to what we want to show, that there is an attainable mixed allocation $\left\{\left(\rho^j, x_i^j, y_i^j\right)\right\}$ that makes some ex ante types better off and none worse off. If type (i,j) is (strictly) better off under the new allocation then clearly

$$\sum_{\eta} q(\eta) \sum_{\theta_i} \lambda(\theta_i,\eta) p(\eta) \cdot x_i^j(\theta_i,\eta) > \sum_{\eta} q(\eta) p(\eta) \cdot \left(y_i^j, (\mu_i - y_i^j) R(\eta)\right),$$

or (x_i^j, y_i^j) would have been chosen. The next thing to show is that for every ex ante type (i,j)

$$\sum_{\eta} q(\eta) \sum_{\theta_i} \lambda(\theta_i,\eta) p(\eta) \cdot x_i^j(\theta_i,\eta) \geq \sum_{\eta} q(\eta) p(\eta) \cdot \left(y_i^j, (\mu_i - y_i^j) R(\eta)\right).$$

If not, then by Assumption 4 it is possible to find a feasible mechanism x_i' such that (x_i', y_i^j) satisfies the budget constraint and makes type i (strictly) better off than (x_i, y_i), contradicting the definition of equilibrium. Then summing the budget constraints for all types (i,j) we have

$$\sum_i \sum_j \rho^j \sum_{\eta} q(\eta) \sum_{\theta_i} \lambda(\theta_i,\eta) p(\eta) \cdot x_i^j(\theta_i,\eta) > \sum_i \sum_j \rho^j \sum_{\eta} q(\eta) p(\eta) \cdot \left(y_i^j, (\mu_i - y_i^j) R(\eta)\right),$$

or

$$\sum_{\eta} q(\eta) p(\eta) \cdot \sum_i \sum_{\theta_i} \lambda(\theta_i,\eta) \sum_j \rho^j x_i^j(\theta_i,\eta) > \sum_{\eta} q(\eta) p(\eta) \cdot \sum_i \sum_j \rho^j \left(y_i^j, (\mu_i - y_i^j) R(\eta)\right). \tag{16}$$

The market-clearing conditions imply that, for each state η, either

$$\sum_i \sum_{\theta_i} \lambda(\theta_i,\eta) \sum_j \rho^j x_i^j(\theta_i,\eta) = \sum_i \sum_j \rho^j \left(y_i^j, (\mu_i - y_i^j) R(\eta)\right)$$

if all the good is consumed at date 1, or $p_1(\eta) = p_2(\eta)$ and

$$\sum_i \sum_{\theta_i} \lambda(\theta_i,\eta) \sum_j \rho^j (x_{i1}^j(\theta_i,\eta) + x_{i2}^j(\theta_i,\eta)) = \sum_i \sum_j \rho^j \left(y_i^j + (\mu_i - y_i^j) R(\eta)\right),$$

if some of the good is stored until date 2. In either case,

$$p(\eta) \cdot \sum_i \sum_{\theta_i} \lambda(\theta_i,\eta) \sum_j \rho^j x^j(\theta_i,\eta) = p(\eta) \cdot \sum_i \sum_j \rho^j \left(y_i^j, (\mu_i - y_i^j) R(\eta)\right).$$

Multiplying this equation by $q(\eta)$ and summing over η yields

$$\sum_{\eta} q(\eta) p(\eta) \cdot \sum_i \sum_{\theta_i} \lambda(\theta_i,\eta) \sum_j \rho^j x^j(\theta_i,\eta) = \sum_{\eta} q(\eta) p(\eta) \cdot \sum_i \sum_j \rho^j \left(y_i^j, (\mu_i - y_i^j) R(\eta)\right),$$

in contradiction of the inequality (**??**). This completes the proof.

8.2 Proof of Proposition 7

Uniqueness. At any price p close to p^*, the short asset is dominated between dates 1 and 2 so no one will hold it. It follows that all consumption at date 1 is provided by the short asset and all consumption at date 2 is provided by the long asset. Then the market-clearing conditions are

$$\begin{array}{l} \sum_{k=1}^{K} \lambda_k \sigma_k x_1(\sigma_k) = ny, \\ \sum_{k=1}^{K} \lambda_k (1-\sigma_k) x_2(\sigma_k) = nr(1-y). \end{array}$$

The budget constraints are

$$\sigma_k x_1(\sigma_k) + (1-\sigma_k) p x_2(\sigma_k) = y + pr(1-y),$$

for each k. For the given (p, y), strict concavity implies that there is at most one solution to the maximization problem. Hence, the equilibrium values are uniquely determined. If there exist two, distinct, regulated equilibria (p, x, y) and (p, x', y'), say, then from the budget constraints and the fact that consumption at each date is a normal good, consumption must be uniformly higher at each date in one regulated equilibrium. But this is clearly impossible from the market-clearing conditions. Hence, there is at most one regulated equilibrium

Existence. Similar to the proof of Proposition 6. ■

8.3 Proof of Proposition 8

Let $X(\sigma_k, p, y) = (X_1(\sigma_k, p, y), X_2(\sigma_k, p, y))$ denote the optimal mechanism for each value of (p, y). An optimal mechanism is one that solves the problem:

$$\begin{array}{ll} \max & \sigma_k U(x_1(\sigma_k)) + (1-\sigma_k) U(x_2(\sigma_k)) \\ \text{s.t.} & \sigma_k x_1(\sigma_k) + (1-\sigma_k) x_2(\sigma_k)/r \leq y + pr(1-y) \end{array}$$

for each k. The necessary and sufficient conditions for this are

$$U'(X_1(\sigma_k; p, y)) = rU'(X_2(\sigma_k; p, y)), \tag{17}$$

and

$$\sigma_k X_1(\sigma_k; p, y) + (1-\sigma_k) p X_2(\sigma_k; p, y) = y + pr(1-y), \tag{18}$$

for every k.

The function $X(\sigma_k, p, y)$ is well defined and continuously differentiable for every (p, y) in a sufficiently small neighborhood of (p^*, y^*). This follows from the implicit function theorem and the regularity of the system (**??**)-(**??**) at (p^*, y^*):

$$\det \begin{bmatrix} U''(X_1(\sigma_k; p^*, y^*)) & -rU''(X_2(\sigma_k; p^*, y^*)) \\ \sigma_k & (1-\sigma_k)p^* \end{bmatrix} \neq 0.$$

The market-clearing condition can be written as

$$\sum_{k=1}^{K} \lambda_k \sigma_k X_1(\sigma_k; p, y) = y,$$

since the other market-clearing condition is automatically satisfied by Walras' law. To show the existence of a regulated equilibrium for each p sufficiently close to p^* we can use the implicit function theorem again, noting that

$$\begin{aligned} \frac{\partial}{\partial y}\left(\sum_{k=1}^{K} \lambda_k \sigma_k X_1(\sigma_k; p, y) - y\right) &= \sum_{k=1}^{K} \lambda_k \sigma_k \frac{\partial X_1(\sigma_k; p^*, y^*)}{\partial y} - 1 \\ &= -1, \end{aligned}$$

because $y + p^* r(1-y) = 1$ is independent of y.

We next calculate the change in expected utility caused by a small change in p at p^*.

$$\begin{aligned} & \frac{d}{dp}\sum_k \lambda_k \left\{\sigma_k U(X_1(\sigma_k; p, y)) + (1-\sigma_k)U(X_2(\sigma_k; p, y))\right\} \\ = & \sum_k \lambda_k \left\{\sigma_k U'(X_1(\sigma_k; p, y))\frac{\partial X_1}{\partial p}(\sigma_k; p, y) + (1-\sigma_k)U'(X_2(\sigma_k; p, y))\frac{\partial X_2}{\partial p}(\sigma_k; p, y)\right\} \\ = & \sum_k \lambda_k U'(X_2(\sigma_k; p, y))\left\{r\sigma_k \frac{\partial X_1}{\partial p}(\sigma_k; p, y) + (1-\sigma_k)\frac{\partial X_2}{\partial p}(\sigma_k; p, y)\right\} \\ = & \sum_k \lambda_k U'(X_2(\sigma_k; p, y))\left\{r(1-y) - (1-\sigma_k)X_2(\sigma_k; p, y)\right\}. \end{aligned}$$

Now suppose that the degree of relative risk aversion is less than one (the other case can be dealt with in exactly the same way). Then (**??**) implies that

$$rx_1(\sigma_k) < rx_2(\sigma_k), \forall \sigma_k.$$

This in turn implies that $x_1(\sigma_k)$ and $(1-\sigma_k)x_2(\sigma_k)$ move in opposite directions. To see this, note that, within a fixed equilibrium, an increase in $(1-\sigma_k)$ must lead to a decrease in consumption at both dates (this follows directly from the fact that $x_1(\sigma_k) < x_2(\sigma_k)/r$). Then $\sigma_k x_1(\sigma_k)$ falls as $(1-\sigma_k)$ increases and so in order to satisfy the budget constraint it must be the case that $(1-\sigma_k)x_2(\sigma_k)/r$ increases. Assuming the distributions are not degenerate, it must be the case that

$$\begin{aligned}
& \sum_k \lambda_k U'(X_2(\sigma_k;p,y))\left\{r(1-y)-(1-\sigma_k)X_2(\sigma_k;p,y)\right\} \\
> & \sum_k \lambda_k U'(X_2(\sigma_k;p,y)) \sum_k \lambda_k \left\{r(1-y)-(1-\sigma_k)X_2(\sigma_k;p,y)\right\} \\
= & \; 0,
\end{aligned}$$

where the last equation follows from the market-clearing condition

$$\sum_k \lambda_k (1-\sigma_k) X_2(\sigma_k;p,y) = r(1-y).$$

Table 1
Examples

Example	r	y	$(q(1), p_2(1))$	$(q(2), p_2(2))$	$(c_1(1), c_2(1))$	$(c_1(2), c_2(2))$	$E[U]$
1A	1.5	0.5	$(0.5, 0.67)$	$(0.5, 0.5)$	$(1, 1.5)$	$(1, 2)$	0.27
1B	0.5	0.64	$(0.61, 1)$	$(0.39, 0.89)$	$(0.82, 0.82)$	$(1, 2)$	0.05
1C	0.3	0.78	$(0.59, 1)$	$(0.41, 1)$	$(0.85, 0.85)$	$(1.21, 1.21)$	0.02
2 A :	0.5	0.5	$(0.5, 0.8)$	$(0.5, 0.8)$	$(1, 1.25)$	$(1, 1.25)$	0.11
2 N :		1.0			$(0, 0.25)$	$(0, 4.75)$	1.25
3A	1.5	0.5	$(0.33 - 0.75,$ $0.25 - 1)$	$(0.25 - 0.67,$ $0.38 - 1)$	$(1, 1.5)$	$(1, 2)$	0.27
3B	0.95	0.49	$(0.55 - 0.77,$ $0.70 - 0.98)$	$(0.45 - 0.33,$ $0.53 - 0.73)$	$(0.97, 0.97)$	$(0.97, 2.05)$	0.16
3C S :	0.7	0.58	$(0.67, 0.88)$	$(0.33, 0.88)$	$(0.94, 0.94)$	$(0.94, 1.7)$	0.09
3C B :		0			$(0.75, 1.5)$	$(0.75, 1.7)$	0.09
4	1.6	0.5	$(0.4, 0.94)$	$(0.6, 0.42)$	$(1.25, 1.33)$	$(0.83, 2)$	17

References

Aghion, P., P. Bacchetta and A. Banerjee (2000). "Currency Crises and Monetary Policy in an Economy with Credit Constraints," working paper, University College, London.

Allen, F. and D. Gale (1994). "Liquidity Preference, Market Participation and Asset Price Volatility," *American Economic Review* **84,** 933-955.

— and — (1998). "Optimal Financial Crises," *Journal of Finance* **53,** 1245-1284.

— and — (1999). *Comparing Financial Systems.* Cambridge: MIT Press.

— and — (2000a). "Financial Contagion," *Journal of Political Economy* **108,** 1-33.

— and — (2000b). "Optimal Currency Crises" *Journal of Monetary Economics*, Carnegie-Rochester Series on Public Policy, to appear.

— and — (2000c). "Banking and Markets Addendum: Prudential Regulation of Banks," www.econ.nyu.edu/user/galed/papers.

Bhattacharya, S. and D. Gale (1987). "Preference Shocks, Liquidity and Central Bank Policy," in William Barnett and Kenneth Singleton, eds., *New Approaches to Monetary Economics* (Cambridge University Press, New York, NY).

Bisin and Gottardi (2000). "Decentralizing Incentive-Compatible Allocations in Adverse Selection Economies," New York University, unpublished.

Bryant, J. (1980). "A Model of Reserves, Bank Runs, and Deposit Insurance," *Journal of Banking and Finance* **4,** 335-344.

Bordo, M. (2000). "The Globalization of International Financial Markets: What Can History Teach Us?" working paper, Rutgers University.

— and B. Eichengreen (2000). "Is the Crisis Problem Growing More Severe?" working paper, Rutgers University.

Calvo, G. (1988). "Servicing the Public Debt: The Role of Expectations," *American Economic Review* **78**, 1411-1428.

Champ, Bruce, Bruce Smith, and Stephen Williamson, 1996. "Currency Elasticity and Banking Panics: Theory and Evidence," *Canadian Journal of Economics* **29**, 828-864.

Chang, R. and A. Velasco (1998a). "Financial Fragility and the Exchange Rate Regime," working paper, New York University.

— and — (1998b). "Financial Crises in Emerging Markets: A Canonical Model," working paper, New York University.

Chari, V. (1989). "Banking Without Deposit Insurance or Bank Panics: Lessons From a Model of the U.S. National Banking System," *Federal Reserve Bank of Minneapolis Quarterly Review* **13** (Summer), 3-19.

— and R. Jagannathan (1988). "Banking Panics, Information, and Rational Expectations Equilibrium," *Journal of Finance* **43**, 749-60.

— and P. Kehoe (2000). "Financial Crises as Herds," working paper, Federal Reserve Bank of Minneapolis.

Cole, H. and T. Kehoe (1995). "Self-fulfilling Debt Crises," *Journal of International Economics* (forthcoming).

Cone, K. (1983). *Regulation of Depository Institutions.* Stanford University Ph.D. dissertation.

Corsetti, G., P. Pesenti and N. Roubini (1999). "Paper Tigers? A Model of the Asian Crisis," *European Economic Review* **43**, 1211-1236.

Diamond, D. (1997). "Liquidity, Banks and Markets," *Journal of Political Economy* 105, 928-956.

— and P. Dybvig (1983). "Bank Runs, Liquidity, and Deposit Insurance," *Journal of Political Economy* **91**, 401-419.

— and R. Rajan (2000). "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking," working paper, University of Chicago.

Flood, R. and N. Marion (1999). "Perspectives on Recent Currency Crises Literature," *International Journal of Finance and Economics* **4**, 1-26.

Geanakoplos, J. and H. Polemarchakis (1986). "Existence, Regularity, and Constrained Suboptimality of Competitive Allocations When the Asset Market Is Incomplete," in Walter, Heller, Ross Starr, and David Starrett eds. *Essays in honor of Kenneth J. Arrow: Volume 3, Uncertainty, information, and communication.* Cambridge, New York and Sydney: Cambridge University Press, 65-95. Also reprinted in Debreu, Gerard, ed. (1996). *General equilibrium theory.* Volume 2. Elgar Reference Collection. International Library of Critical Writings in Economics, no. 67. Cheltenham, U.K.: Elgar; distributed by Ashgate, Brookfield, Vt. 67-97.

Hellwig, M. (1994). "Liquidity Provision, Banking, a nd the Allocation of Interest Rate Risk," *European Economic Review* **38**, 1363-1389.

— (1998). "Banks, Markets, and the Allocation of Risks in an Economy," *Journal of Institutional and Theoretical Economics* **154**, 328-345.

Jacklin, C. (1986). "Demand Deposits, Trading Restrictions, and Risk Sharing," in E. Prescott and N. Wallace, eds., *Contractual Arrangements for Intertemporal Trade* (University of Minnesota Press, Minneapolis, MN).

— and S. Bhattacharya (1988). "Distinguishing panics and information-based bank runs: Welfare and policy implications," *Journal of Political Economy* **96**, 568-592.

Kaminsky, G. and C. Reinhart (1999). "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems," *American Economic Review* **89**, 473-500.

Krugman, P. (1979). "A Model of Balance-of-Payments Crises," *Journal of Money Credit and Banking* **11**, 311-325.

Morris, S. (2000). "Contagion," *Review of Economic Studies* **67**, 57-78.

Morris, S. and H. S. Shin (1998). "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *American Economic Review* **88**, 587-597.

Morris, S. and H. S. Shin (2000). "Global Games: Theory and Applications," working paper, Yale University.

Obstfeld, M. (1986). "Rational and Self-Fulfilling Balance of Payments Crises," *American Economic Review* **76**, 72-81.

Peck, J. and K. Shell (1999). "Bank Portfolio Restrictions and Equilibrium Bank Runs," working paper, Cornell University.

Postlewaite, A. and X. Vives (1987). "Bank Runs as an Equilibrium Phenomenon," *Journal of Political Economy* **95**, 485-491.

Melitz, J. (1990). "Financial Deregulation in France," *European Economic Review* **34**, 394-402.

Prescott, E. and R. Townsend (1984a). "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard, *Econometrica* **52**, 21-45.

Prescott, E. and R. Townsend (1984b). "General Competitive Analysis in an Economy with Private Information," *International Economic Review* **25**, 1-20.

Wallace, N. (1988). "Another Attempt to Explain an Illiquid Banking System: The Diamond and Dybvig Model with Sequential Service Taken Seriously," *Federal Reserve Bank of Minneapolis Quarterly Review* **12** **(**Fall), 3-16.

Wallace, N., (1990). "A Banking Model in which Partial Suspension is Best," *Federal Reserve Bank of Minneapolis Quarterly Review* **14** (Fall), 11-23.