We study the interaction between a government’s bailout policy during a banking crisis and individual banks’ willingness to impose losses on (or “bail in”) their investors. We consider an environment in which banks and investors are free to write complete, state-contingent contracts. Our primary focus is on the timing of this contract’s response to an incipient crisis. In the constrained efficient allocation, banks facing losses immediately cut payments to withdrawing investors. In a competitive equilibrium, however, these banks often delay cutting payments in order to benefit more from the eventual bailout. In some cases, the costs associated with this delay are large enough that investors will choose to run on their bank, creating further distortions and deepening the crisis. Our approach has novel implications for the form a banking crisis must take. For example, a bank run cannot be driven purely by sunspots in our model; it can only occur at banks that have suffered some real losses. In addition, a run can only occur when these losses are systemic, that is, experienced by a large number of banks at once. This run can nevertheless be self-fulfilling in the sense that investors run when their bank suffers losses if and only if they expect other investors to do the same. We discuss the implications of the model for banking regulation and optimal policy design.

1 Introduction

In the years since the financial crisis of 2008 and the associated bailouts of financial institutions, policy makers in several jurisdictions have drafted rules requiring that these institutions impose losses on (or “bail in”) their investors in any future crisis. These rules aim both to protect taxpayers in the event of a future crisis and to change the incentives of banks and investors.

---

*PRELIMINARY AND INCOMPLETE DRAFT. We thank seminar participants at Rutgers, the 2016 Oxford Financial Intermediation Theory Conference, and especially Joel Shapiro for helpful comments

†Rutgers University: todd.keister@rutgers.edu

‡Rutgers University: ymitkov@econ.rutgers.edu
in a way that makes such a crisis less likely. While the specific requirements vary, and are often yet to be finalized, in many cases the bail-in will be triggered by an announcement or action taken by the institution facing losses. This fact raises the question of what incentives banks will face when deciding whether and when to bail in their investors. In this paper, we study the interaction between a government’s bailout policy and individual banks’ willingness to take actions that bail in their investors. In particular, we ask how the prospect of being bailed out by the government changes bank’s bail-in decisions and how these decisions, in turn, affect the susceptibility of the banking system to a run by investors.

At one level, the reason why banks and other financial intermediaries sometimes experience runs by their investors is well understood. Banks offer deposit contracts that allow investors to withdraw their funds at face value on demand or at very short notice. During a financial panic, investors fear that a combination of real losses and/or heavy withdrawals will make their bank insolvent and therefore unable to meet all of its obligations. This belief makes it individually rational for each investor to withdraw her funds at the first opportunity; the ensuing rush to withdraw then guarantees that the bank does indeed fail, justifying investors’ pessimistic beliefs.¹

A key element of this well-known story is that the response to a bank’s losses and/or a run by its investors is delayed. In other words, there is a period of time during which a problem clearly exists and investors are rushing to withdraw, but the bank continues to operate as normal. Only when the situation becomes bad enough is some action – freezing deposits, renegotiating obligations, imposing losses on some investors, etc. – taken. This delay tends to deepen the crisis and thereby increase the incentive for investors to withdraw their funds at the earliest opportunity.

From a theoretical perspective, this delayed response to a crisis presents something of a puzzle. A run on the bank creates a misallocation of resources that makes the bank’s investors as a group worse off. Why do these investors not collectively agree to an alternative arrangement that efficiently allocates whatever losses have occurred while minimizing liquidation and other costs? In particular, why does the banking arrangement not respond more quickly to whatever news leads investors to begin to panic and withdraw their funds?

Most of the literature on bank runs resolves this puzzle using an incomplete-contracts approach. In particular, it is typically assumed to be impossible to write and/or enforce the type of contracts that would be needed to generate fully state-contingent payments to investors. The classic paper of Diamond and Dybvig (1983), for example, assumes that banks must pay withdrawing investors at face value until the bank has liquidated all of its assets and is completely out of funds. Other contracts – in which, for example, the bank is allowed to impose withdrawal fees when facing a run – are simply not allowed. Even those more recent papers

¹This basic logic applies not only to commercial banking to also to a wide range of financial intermediation arrangements. See Yorulmazer (2014) for a discussion of a several distinct financial intermediation arrangements that experienced run-like episodes during the financial crisis of 2008.
that study optimal banking arrangements, and that take into account the possibility of a run, impose some incompleteness of contracts. Peck and Shell (2003), for example, allow a bank to adjust payments to withdrawing investors based on any information it receives. However, the bank is assumed not to observe the realization of a sunspot variable that is available to investors and, in this sense, the ability to make state-contingent payments is still incomplete.2

If the fundamental problem underlying the fragility of banking arrangements is incompleteness of contracts, then a primary focus of policy makers should be on removing this incompleteness. In other words, a central policy conclusion of the literature to date is that financial stability policy should aim to create legal structures under which more fully state-contingent banking contracts become feasible. There has, in fact, been substantial progress in this direction in recent years, including the establishment of orderly resolution mechanisms for large financial institutions and other ways of “bailing in” these institutions’ investors more quickly and more fully than in the past. The reform of money market mutual funds that was adopted in the U.S. in 2014 is a prime example. Under the new rules, these funds are permitted to temporarily prohibit redemptions (called “erecting a gate”) and impose withdrawal fees during periods of high withdrawal demand if doing so is deemed to be in the best interests of the funds’ investors.

In this paper, we ask whether making banking arrangements more fully state contingent – thereby allowing banks increased flexibility to bail in their investors – is sufficient to eliminate the problem of bank runs. To answer this question, we study a model in the tradition of Diamond and Dybvig (1983), but in which banks can freely adjust payments to investors based on any information available to the bank or to its investors. We think of this assumption as capturing an idealized situation in which policy makers’ efforts to improve the contractual environment have been completely successful.

There are two aggregate states in our model and banks face uncertainty about the value of their investments. No banks experience losses in the good aggregate state, but in the bad aggregate state, some banks’ assets are impaired. The government is benevolent and taxes agents’ endowments in order to provide a public good. If there is a banking crisis, the government can also use these resources to provide bailouts to impaired banks. The government observes the aggregate state but cannot immediately tell which banks have impaired assets and which do not. In addition, the government cannot commit to a bailout plan; instead, the payment made to each bank will be chosen as a best response to the situation at hand. As in Keister (2016), this inability to commit implies that banks in worse financial conditions will receive larger bailout payments, as the government will choose to equalize the marginal utility of consumption across all depositors to the extent possible.

2This same approach is taken in a large number of papers that study sunspot-driven bank runs in environments with flexible banking contracts, including Ennis and Keister (2010a), Sultanum (2015), Keister (2016), and many others. See Andolfatto et al. (2016) for an interesting model in which the bank does not observe the sunspot state, but can attempt to elicit this information from investors.
A bank with impaired assets has fewer resources available to make payments to investors. In an efficient allocation, such a bank would respond by immediately bailing in its investors, reducing all payments so that the loss is evenly shared. When the bank anticipates a government intervention, however, it may have an incentive to delay this response. By instead acting as if its assets were not impaired, the first group of its depositors who withdraw will receive higher payments. The government will eventually learn that the bank’s assets are impaired and, at this point, will find the bank to be in worse financial shape as a result of the delayed response. The inability to commit prevents the government from being able to punish the bank at this point; instead, the bank will be given a larger bailout payment as the government aims to raise the consumption levels of its remaining investors. This larger payment then justifies the bank’s original decision to delay taking action. In other words, we show that bailouts delay bail-ins.

The delay in banks’ bail-in decisions has implications at both the aggregate and bank level. The delayed response makes banks with weak fundamentals even worse off and leads the government to make larger bailout payments, at the cost of a lower level of public good provision for everyone. In some cases, the misallocation of resources created by the delay may be large enough to give investors in weak banks an incentive to run in an attempt to withdraw before the bail-in is enacted. In these cases, the delayed bail-in creates financial fragility.

Our approach has novel implications for the form a banking crisis must take. Models in the tradition of Diamond and Dybvig (1983) typically do not distinguish between a single bank and the banking system; one can often think of the same model as applying equally well to both situations. If the banking system is composed of many banks, such models predict that there could be a run on a single bank, on a group of banks, or on all banks, depending on how each bank’s depositors form their beliefs. In our model, in contrast, there cannot be a run on only one bank, nor can there be a crisis in which only one bank chooses to delay bailing in its investors. If there is only a problem at one bank in our model, the government will choose to provide full deposit insurance, which removes any incentive for investors to run as well as any need for the bank to enact a bail-in. Individual banks can fail in this model because the return on their assets is random, but the problems we focus on here (bank runs and delayed bail-ins) can only arise if the underlying losses are sufficiently widespread.

In addition, a bank run in this model cannot be driven purely by sunspots. If a bank’s assets are not impaired, we show that there is no incentive for it to delay its response should a run occur. Given that there is no delay in the response, the bank’s depositors will have no incentive to run. The model also displays a non-linear relationship between the aggregate losses on banks’ investments and the resulting social costs. When the number of banks experiencing a loss is small enough, there is no delay in equilibrium and the only losses come from the low realization of investment returns in some banks. As the total losses increase, however, eventually it becomes optimal for impaired banks to delay their response. This fact leads to a misallocation of resources that amplifies the effects of the underlying shock.
The remainder of the paper is organized as follows. The next section describes the economic environment and the elements of our model, including the strategies available to banks, investors, and the government. In Section 3, we present the efficient allocation of resources in this environment, which is a useful benchmark for what follows. Section 4 contains the heart of the analysis: a discussion of best responses and the construction of an equilibrium with a bank run and delayed response. We explore some further implications of the analysis in Section 5.

2 The model

We analyze a version of the Diamond and Dybvig (1983) model with an explicit sequential service constraint and fiscal policy conducted by a government with limited commitment. We introduce idiosyncratic risk to banks’ asset holdings in this environment and highlight how bank’s incentives to react to losses are influenced by their anticipation of government intervention.

2.1 The environment

There are three time periods, labeled $t = 0, 1, 2$. In the paragraphs that follow, we introduce the agents, preferences, and technologies that characterize the economic environment.

**Investors.** There is a continuum of investors, indexed by $i \in [0, 1]$, in each of a continuum of locations, indexed by $k \in [0, 1]$. Each investor has preferences characterized by

$$U(i_{1}^{i,k}, i_{2}^{i,k}, g; \omega_{i,k}) \equiv u(c_{1}^{i,k} + \omega_{i,k}c_{2}^{i,k}) + v(g),$$

(1)

where $c_{t}^{i,k}$ denotes the period-$t$ private consumption of investor $i$ in location $k$ and $g$ is the level of the public good, which is available in all locations. The random variable $\omega_{i,k} \in \Omega \equiv \{0, 1\}$ is realized at $t = 1$ and is privately observed by the investor. If $\omega_{i,k} = 0$, she is impatient and values private consumption only in period 1, whereas if $\omega_{i,k} = 1$ she values consumption equally in both periods. Each investor will be impatient with a known probability $\pi > 0$, and the fraction of investors who are impatient in each location will also equal $\pi$. The functions $u$ and $v$ are assumed to be smooth, strictly increasing, strictly concave and to satisfy the usual Inada conditions. As in Diamond and Dybvig (1983), the function $u$ is assumed to exhibit a coefficient of relative risk aversion that is greater than one for all $c > 0$. Each investor is endowed with one unit of of an all-purpose good at the beginning of period 1 and nothing in subsequent periods. Investors cannot directly invest their endowments and must instead deposit with a financial intermediary.
**Banks.** In each location, there is a representative financial intermediary that we refer to as a bank. Each bank accepts deposits in period 0 from investors in its location and allows these investors to withdraw in either period 1 or period 2. Investors are isolated from each other at all times and those who choose to withdraw in period 1 arrive at their bank sequentially in the order determined by their index $i$. We assume that investors do not know this order in period 0 and, therefore, are ex ante identical. At the start of period 1, before withdrawal decisions are made, each investor privately observes both her preference type $\omega_{i,k}$ and her position in the withdrawal decision order. At this point, investor $(i,k)$ knows that if she chooses to withdraw early, she will arrive at bank $k$ before all investors $(i',k)$ with index $i' > i$. When an investor arrives, the bank chooses how much to pay her and the investor must consume these goods immediately upon receiving them. These features of the environment give rise to a sequential service constraint, as in Wallace (1988) and others, where the consumption of each investor can only be contingent on the information available to the bank at the time of her withdrawal.

Each bank invests the deposits it receives at $t = 0$ in a set of ex ante identical projects. A project requires one unit of input at $t = 0$ and offers a gross return of 1 at $t = 1$ or of $R > 1$ at $t = 2$ if it is not impaired. In period 1, however, a fraction $\sigma_k \in \Sigma = \{0, \bar{\sigma}\}$ of the projects held by bank $k$ will be revealed to be impaired. An impaired project is worthless: it produces zero return in either period. We will refer to $\sigma_k$ as the fundamental shock to bank $k$. If $\sigma_k = 0$, the bank is said to have sound fundamentals, whereas if $\sigma_k = \bar{\sigma} > 0$, the bank has weak fundamentals. Investors observe the value of $\sigma_k$ for their bank at $t = 1$ and can base their withdrawal decision on this information.

At $t = 0$, each bank chooses a contract that specifies the payment it will give to each withdrawing depositor in each state of nature. In particular, this contract specifies, for each possible value of $\sigma_k$, how much consumption the first investor to withdraw receives, how much the second investor to withdraw receives, and so on. Note that each payment can depend on the number of withdrawals that have previously occurred, but not on the withdrawing investor’s preference types because that information is not available to the bank. The contract is chosen to maximize the expected utility of the banks’ investors, and the bank is committed to follow this contract. As in Wallace (1988), a bank here can be thought of as an automated teller machine (ATM) that is programmed by its investors at $t = 0$ and simply makes payments according to this program in later periods. In taking this approach, we aim to capture a contractual environment that is sufficiently rich to eliminate any agency problems between bankers and their investors.

---

3 Even though we use the term “bank” for simplicity, our model should be interpreted as applying to a wide range of financial institutions that engage in maturity transformation.

4 Our formulation of the sequential service constrained follows Green and Lin (2000, 2003) and Ennis and Keister (2009b, 2010a) in assuming that the investors know their precise position in the withdrawal decision order. We could alternatively assume that investors decide when to withdraw before knowing this position, but can change their mind if they arrive at their bank after the government has intervened. The same results obtain under this alternative formulation, but the notation required is slightly more complex.
Aggregate uncertainty. The fraction of banks whose assets are impaired depends on the aggregate state of the economy, which is either good or bad. In the good state, $\sigma_k = 0$ for all banks, whereas in the bad state $\sigma_k = \bar{\sigma}$ for a fraction $n \in [0, 1]$ of banks. Total losses in the financial system in the bad state thus equal $n\bar{\sigma}$. The probability of the bad state is denoted $q$; we interpret this event as an economic downturn that has differing effects across banks. If we think of the projects in the model as representing loans, for example, then the loans made by some banks are relatively unaffected by the downturn (for simplicity, we assume they are not affected at all), while other banks find they have substantial non-performing loans. Conditional on the bad aggregate state, all banks are assumed equally likely to experience weak fundamentals. The ex-ante probability that a given bank’s fundamentals will be weak is, therefore, equal to $qn$.

The government. The government in our model acts as both a fiscal authority and a banking supervisor. Its objective is to maximize the sum of all investors’ expected utilities at all times. The government’s only opportunity to raise revenue comes in period 0, when it chooses to tax investors’ endowments at rate $\tau$. In period 1, the government will use this revenue to provide the public good and, perhaps, to make transfers (bailouts) to banks. The government is unable to commit to the details of the bailout intervention ex-ante, but instead chooses the policy ex post, after it learns the full state and some payments have already been made by banks.

The government observes the aggregate state at the beginning of period 1, but is initially unable to determine which banks have experienced losses when the state is bad. After a measure $\theta \geq 0$ of investors have withdrawn from each bank, the government observes the idiosyncratic state $\sigma_k$ of all banks and decides how to allocate its tax revenue between bailout payments to banks with weak fundamentals and the public good. Banks that receive a bailout from the government are immediately placed in resolution and all subsequent payments made by these banks are chosen by the government. The parameter $\theta$ thus measures how quickly the government can collect bank-specific information during a crisis and respond to this information. Once the public good has been provided, the government has no access to resources and, hence, there will be no further bailouts.

2.2 Timeline

The sequence of events is depicted on Figure (1). Period 0 starts with the government taxing all endowments at rate $\tau$ and then investors depositing their after-tax endowments in the

---

5Alternatively, we could allow for the banking contract set at $t = 0$ to specify these payments as a function of the size of the bailout payment received at $t = 1$. The two approaches lead to exactly the same results; having the government mandate the payments simplifies the presentation.
banking sector. At the start of period 1, each investor observes (privately) whether she is patient or impatient. In addition, investors and their bank (but not the government) observe the bank’s fundamental state. Investors then make their withdrawal decisions and those choosing to withdraw early begin arriving at their banks. Once the measure of withdrawals reaches $\theta$, the government observes all banks’ idiosyncratic states, bailout payments are made, and all remaining tax revenue is used to provide the public good. Banks then continue to make payments to investors choosing to withdraw in period 1. In period 2, all remaining investors withdraw and the game ends.

2.3 Discussion

[To be written.]

3 The constrained efficient allocation

We begin by studying an allocation that will serve as a useful benchmark in the analysis. Suppose a benevolent planner could control the operations of all banks and the government, but not investors’ withdrawal decisions. This planner observes all of the information available to banks, but does not observe investors’ preference types. Like the government, the planner is unable to commit to future actions. The planner’s objective at each decision node is to maximize the sum of all investors’ expected utilities. We derive the allocation that would be
implemented by such a planner by working backward through the timeline in Figure 1.6

3.1 Post-bailout payments

We start with point $(d)$ on the timeline, after a fraction $\theta$ of investors have withdrawn from each bank and any bailout payments have been made. We ask how the planner would want the remaining resources in each bank to be allocated. Let $\hat{\rho}_k$ denote the fraction of the remaining investors in bank $k$ who have chosen to withdraw in period 1. Let $\hat{\psi}_k$ denote the per-capita resources of bank $k$, including any bailout payment that it has received. Because investors are risk averse and the planner’s objective is to maximize the sum of their expected utilities, the planner will want all investors in a bank who withdraw in the same period to consume the same amount. Letting $(\hat{c}_1, \hat{c}_2)$ denote these amounts, the planner will want the payments made by bank $k$ to solve

$$\hat{V}(\hat{\psi}_k; \hat{\rho}_k) \equiv \max_{[\hat{c}_1, \hat{c}_2]} \hat{\rho}_k u(\hat{c}_1) + (1 - \hat{\rho}_k) u(\hat{c}_2)$$

subject to the resource constraint

$$\hat{\rho}_k \hat{c}_1 + (1 - \hat{\rho}_k) \frac{\hat{c}_2}{R} \leq \hat{\psi}_k.$$

The first-order condition for this problem is

$$u'(\hat{c}_1) = R u'(\hat{c}_2) = \hat{\mu},$$

where $\hat{\mu}$ denotes the multiplier on the resource constraint. Together, this condition and the resource constraint determine the solution to the problem, which we denote

$$(\hat{c}_1(\hat{\psi}_k; \hat{\rho}_k), \hat{c}_2(\hat{\psi}_k; \hat{\rho}_k)),$$

and we use $\hat{\mu}(\hat{\psi}_k; \hat{\rho}_k)$ to denote the value of the multiplier at this solution. Note that the first-order condition implies that we have

$$\hat{c}_1(\hat{\psi}_k; \hat{\rho}_k) < \hat{c}_2(\hat{\psi}_k; \hat{\rho}_k) \text{ for all } \hat{\psi}_k, \hat{\rho}_k.$$  

In other words, the way in which the planner desires to allocate the remaining resources within bank $k$ in response to any set of withdrawal decisions by its investors has the property that an investor will always receive more if she waits until period 2 than if she withdraws in period 1. As we discuss in more detail below, this condition immediately implies that patient investors will never have an incentive to run on their bank after bailout payments have been made.

6We define the constrained efficient allocation using a game played between the planner and investors, with the latter making withdrawal decisions based on their own privately-observed types. We show below that this game has a unique equilibrium and, hence, the allocation the planner would choose to implement through this game is well defined.
3.2 Efficient bailouts

We now move to point (c) in the timeline in Figure 1, where the planner chooses how to allocate the existing tax revenue between bailout payments and the provision of the public good. In particular, the planner chooses a bailout payment \( b_k \geq 0 \) to give to each bank \( k \) as a function of the current situation. We restrict the planner’s actions to satisfy

\[
 b_k(\sigma_k) = 0 \text{ if } \sigma_k = 0, \tag{5}
\]

that is, banks with strong fundamentals do not receive bailouts. Let \( c_k^1 \) denote the average payment given by bank \( k \) to the first \( \theta \) of its investors to withdraw, meaning that it has paid a total of \( \theta c_k^1 \) to these investors. Then the planner will choose the bailout payments to solve

\[
 \max_{\{b_k \geq 0\}} \int_{k \in L} \hat{V} \left( \frac{(1 - \tau) (1 - \bar{\sigma}) - \theta c_k^1 + b_k}{1 - \theta}; \hat{\rho}_k \right) dk + v \left( \tau - \int_{k \in L} b_k dk \right), \tag{6}
\]

where \( L \) denotes the set \( \{ k : \sigma_k = \bar{\sigma} \} \), that is, the set of banks that have experienced a loss. The first term in this expression measures the utility from private consumption of all remaining investors in banks with weak fundamentals, while the second term captures the fall in public consumption for all investors that occurs when some tax revenue is used for bailouts.

The first-order condition for the choice of \( b_k \) can be written as

\[
 \hat{\mu} \left( \frac{(1 - \tau) (1 - \bar{\sigma}) - \theta c_k^1 + b_k}{1 - \theta}; \hat{\rho}_k \right) \leq v' \left( \tau - \int_{k \in L} b_k dk \right) \text{ for all } k \in L, \tag{7}
\]

and this condition must hold with equality for all banks with \( b_k > 0 \). To understand what this condition implies, imagine that banks have differing levels of resources, perhaps because they have chosen different levels of \( c_k^1 \). The planner will first allocate bailout payments to the banks with the fewest remaining resources, whose investors are facing the lowest levels of private consumption. As these bailout payments raise the consumption of these investors, banks in slightly better condition will begin to receive bailout payments, and this process will continue until the marginal utility of private consumption for investors in all banks receiving bailouts equals the marginal utility of public consumption. Notice that equation (7) implies (i) the consumption plan \( (\hat{c}_1^k, \hat{c}_2^k) \) will be the same in all banks receiving a bailout and (ii) any two banks that have the same average early payment \( c_i^k \) will receive the same bailout payment \( b_k \).
3.3 Efficient banking contracts

We now ask how the planner will set banks’ payment contracts. For each bank, the contract specifies the payment to be made to each of the first $\theta$ investors to withdraw at $t = 1$ (see the point labeled $(b)$ in Figure 1.) After this point, the contract specifies the payments made to remaining investors if the bank is not bailed out. For banks that are bailed out, the contract is discarded when the bailout payment is made and the planner will dictate the payments made to the remaining investors by solving problem (2) above. To begin, note that since investors are risk averse, the planner will choose to give the same payment to each of the first $\theta$ investors to withdraw within the same bank. In other words, the value of $c^k_1$ that was defined above as the average payment to these investors in bank $k$ will also be the actual payment given to each of them. This payment will depend on the anticipated withdrawal decisions of the bank’s investors; let $\rho_k$ denote the fraction of all of bank $k$’s investors that the planner forecasts will withdraw in period 1. The first $\theta$ of these investors will each receive the payment $c^k_1$, and then the fraction of the remaining investors who will withdraw in period 1 is given by

$$\hat{\rho}_k = \frac{\rho_k - \theta}{1 - \theta}. \tag{8}$$

Next, note that since the planner’s objective is to maximize the equal-weighted sum of all investors’ expected utilities, it will choose the same payment $c^k_1$ for all banks facing the same idiosyncratic state $\sigma_k$ and withdrawal demand $\rho_k$. Within a given aggregate state, therefore, we can reduce the problem to one of choosing a payment function for each possible bank-specific state:

$$c_1(\sigma_k; \rho_k) \text{ for } \sigma_k \in \{0, \bar{\sigma}\} \text{ and } \rho_k \in [\pi, 1].$$

These payment functions will be chosen to maximize the sum of all investors’ expected utilities, which can be written as

$$\int_{k \in L} \left( \theta u \left(c^k_1(0; \rho_k)\right) + (1 - \theta) \hat{V} \left( \frac{(1 - \tau) - \theta c^k_1(0; \rho_k)}{1 - \theta}; \hat{\rho}_k \right) \right) dk$$

$$+ \int_{k \in L} \left( \theta u \left(c^k_1(\bar{\sigma}; \rho_k)\right) + (1 - \theta) \hat{V} \left( \frac{(1 - \tau)(1 - \bar{\sigma}) - \theta c^k_1(\bar{\sigma}; \rho_k) + b_k}{1 - \theta}; \hat{\rho}_k \right) \right) dk$$

$$+ \nu \left( \tau - \int_{k \in L} b_k dk \right) \tag{9}$$

where the bailout payments $b_k$ will be set as functions of the choices $c^k_1$ according to equation

\footnote{Recall that a bank in our model is a technology that accepts inputs at $t = 0$ and pays out consumption to withdrawing investors at $t = 1$ and $t = 2$. While we assume that the planner, like the government, is unable to commit to its own future actions, the bank is committed to a payment schedule once it has been programmed by the planner. It is relatively straightforward to show that the exact same results would obtain in this section if we were to instead assume that the planner chooses each payment as it is made, with no commitment to a banking contract.}
(7) and the post-bailout withdrawal demand \( \hat{\rho}_k \) is as specified in equation (8).\(^8\) Note that the function \( \hat{V} \) from equation (2) appears in two distinct places in this objective function. On the first line, \( \hat{V} \) represents the maximum expected utility the planner can obtain for the remaining \((1 - \theta)\) investors when there is no bailout and the contract set at \( t = 0 \) remains in force. One the second line, in contrast, \( \hat{V} \) represents the maximum expected utility the planner can obtain when the bank is placed into resolution after a bailout payment has been made. We examine the first-order conditions for this problem for each type of bank separately, beginning with banks that have \( \sigma_k = 0 \).

**Banks with sound fundamentals.** The first-order condition for \( c^k_1 (0; \rho_k) \) can be written as

\[
u' \left( c^k_1 (0; \rho_k) \right) = \hat{\mu} \left( \frac{(1 - \tau) - \theta c^k_1 (0; \rho_k)}{1 - \rho_k} \right) \hat{\rho}_k
\]

Using equation (3), this condition implies that the payments made to investors in a bank with sound fundamentals will satisfy

\[
u' \left( c^k_1 (0; \rho_k) \right) = \nu' \left( \hat{c}^k_1 (0; \hat{\rho}_k) \right) = Ru' \left( \hat{c}^k_2 (0; \hat{\rho}_k) \right)
\]

(10)
as well as the bank-\( k \) resource constraint

\[
\theta c^k_1 (0; \rho_k) + (\rho_k - \theta) \hat{c}^k_1 (0; \hat{\rho}_k) + (1 - \rho_k) \frac{\hat{c}^k_2 (0; \hat{\rho}_k)}{R} \leq 1 - \tau.
\]

(11)

This last constraint states that all payments made to bank \( k \)'s investors are financed by the proceeds of the \( 1 - \tau \) investment projects owned by bank \( k \).\(^9\)

Equations (10) and (11) can be solved for the optimal payments \( \left( c^k_1, c^k_1, c^k_2 \right) \) for any given tax rate \( \tau \). Note that this solution will satisfy

\[
c^k_1 (0; \rho_k) = c^k_1 (0; \hat{\rho}_k) < c^k_2 (0; \hat{\rho}_k).
\]

(12)

In other words, the efficient banking contract has the feature that, in a bank with sound fundamentals, all investors who withdraw from bank \( k \) in period 1 will receive the same amount of consumption regardless of where they fall in the order of withdrawals. In addition, investors who withdraw in period 2 will always receive more than investors who withdraw in period 1.

\(^8\)In other words, the payments \( b_k \) are functions of the choice variables \( c^k_1 \) in this problem. This dependence is not explicitly noted in the objective function simply to save space. Also note that, since the payments \( b_k \) are chosen optimally, the envelope conditions imply that the terms involving \( db_k/dc^k_1 \) drop out of the first-order conditions below.

\(^9\)Keep in mind that the planner is constrained: it is not able to freely reallocate resources across banks. Like the government in the decentralized setting, the planner is can only collect tax revenue in period 0 and make non-negative transfers to banks (i.e., bailouts) in period 1. For a bank that does not receive a bailout payment, this constraint implies that the consumption of its investors will exactly equal the proceeds of its asset portfolio, as shown in equation (11).
regardless of the level of withdrawal demand $\rho_k$. This solution also shows that the payments the planner makes to investors in banks with sound fundamentals do not depend on the number of banks with weak fundamentals nor on the financial condition of these banks. We can think of these payments as representing the “face value” of the banking contract when banks are operated by the planner.

**Banks with weak fundamentals.** We now ask how the planner will chose to have a bank’s payments set when $\sigma_k = \bar{\sigma}$. The first-order condition for the choice of $c_1^k (\bar{\sigma}; \rho_k)$ that maximizes (9) can be written as

$$u' \left( c_1^k (\bar{\sigma}; \rho_k) \right) = \hat{\mu} \left( \frac{(1 - \tau)(1 - \bar{\sigma}) - \theta c_1^k (\bar{\sigma}; \rho_k) + b_k}{1 - \theta} ; \hat{\rho}_k \right).$$

Again using equation (3), this condition implies that the payments made to investors in a bank with weak fundamentals will satisfy

$$u' \left( c_1^k (\bar{\sigma}; \rho_k) \right) = u' \left( \hat{c}_1^k (\bar{\sigma}; \hat{\rho}_k) \right) = Ru' \left( c_2^k (\bar{\sigma}; \hat{\rho}_k) \right)$$

as well as the bank-$k$ resource constraint

$$\theta c_1^k (\bar{\sigma}; \rho_k) + (\rho_k - \theta) \hat{c}_1^k (\bar{\sigma}; \hat{\rho}_k) + (1 - \rho_k) \frac{\hat{c}_2^k (\bar{\sigma}; \hat{\rho}_k)}{R} \leq (1 - \tau)(1 - \bar{\sigma}) + b_k.$$  

The bailout payment $b_k$ in this constraint will be chosen to satisfy the first-order condition in (7). Recall that this choice of $b_k$ will depend on aggregate conditions in the economy, including how many banks have weak fundamentals and the level of remaining resources in these banks. This fact implies that, unlike the case of sound banks discussed above, the consumption of investors in weak banks will in general depend on factors outside of their own banks. Regardless of these factors, though, condition (13) implies that the payments made to investors in bank $k$ will satisfy

$$c_1^k (\bar{\sigma}; \rho_k) = \hat{c}_1^k (\bar{\sigma}; \hat{\rho}_k) < \hat{c}_2^k (\bar{\sigma}; \hat{\rho}_k).$$

These relationships show that all investors who withdraw in period 1 receive the same level of consumption from a bank with weak fundamentals, regardless of whether they withdraw before or after bailout payments have been made. In addition, investors who withdraw in period 2 always receive more than investors who withdraw in period 1.

### 3.4 Withdrawal behavior

We do not allow the planner to observe the preference types of individual investors or to control their choice of when to withdraw. Instead, we have assumed that the planner takes the withdrawal decisions of investors in each bank as given and allocates resources as a best
response to these decisions. The properties of this best response, specifically conditions (4), (12) and (15), have a striking implication: an investor in any bank will always receive more consumption if she withdraws in period 2 than if she withdraws in period 1. This relationship holds regardless of how many other investors in her bank (or elsewhere) choose to withdraw early. In other words, if we take the planner’s operation of the banking system as given and look at the game played by investors in choosing when to withdraw, it is a strictly dominant strategy for each investor in this game to withdraw in period 1 if and only if she is impatient.

It follows immediately that there is a unique equilibrium of this game. We define the allocation that obtains in this equilibrium to be the constrained efficient allocation in our environment. It is fairly easy to see that this allocation is the same as what the planner would choose if it were able to observe investors’ preference types and dictate withdrawal decisions. Note that no event resembling a bank run, in which some patient investors withdraw early, occurs here.

### 3.5 Efficient taxation

Finally, we consider the planner’s decision at point \((a)\) in Figure 1, which is the choice of a tax rate \(\tau \geq 0\). This rate is chosen in period 0, before the aggregate state is realized. In making this decision, the planner knows that the good aggregate state will occur with probability \((1 - q)\), in which case all banks will have sound fundamentals, and the bad state will occur with probability \(q\), in which case a fraction \(n\) of banks will have weak fundamentals. The planner also recognizes that only impatient investors will withdraw early and, therefore, the fraction of investors who withdraw in period 1 will be \(\pi\) in all banks. After \(\theta\) withdrawals have occurred, the fraction of the remaining investors who will withdraw in period 1 will equal

\[
\hat{\pi} = \frac{\pi - \theta}{1 - \theta}
\]

in all banks. The analysis above then shows that the planner will choose the early payment \(c_1(0; \pi)\) for all banks with sound fundamentals, regardless of the aggregate state, and this payment will depend on the tax rate \(\tau\) as derived above. Similarly, a common early payment \(c_1(\bar{\sigma}; \pi)\) will be chosen for all banks with weak fundamentals in the bad aggregate state. Using these solutions and the value function defined in (2), we can write the planner’s objective function in choosing the tax rate \(\tau\) as

\[
(1 - q) \left\{ \theta u \left( c_1(0; \pi) \right) + (1 - \theta) \hat{V} \left( \frac{(1 - \tau) - \theta c_1(0; \pi)}{1 - \theta}; \hat{\pi} \right) + v(\tau) \right\}
\]

\[
+q \left\{ (1 - n) \left( \theta u \left( c_1^k(0; \pi) \right) + (1 - \theta) \hat{V} \left( \frac{(1 - \tau) - \theta c_1(0; \pi)}{1 - \theta}; \hat{\pi} \right) \right) + n \left( \theta u \left( c_1(\bar{\sigma}; \pi) \right) + (1 - \theta) \hat{V} \left( \frac{(1 - \tau)(1 - \bar{\sigma}) - \theta c_1(\bar{\sigma}; \pi)}{1 - \theta}; \hat{\pi} \right) + v(\tau - nb) \right\}
\]

14
The first line of this expression represents the expected utility of all investors in the good aggregate state, when all banks are sound. The second line represents the expected utility from private consumption of the fraction \((1 - n)\) of investors who, in the bad aggregate state, find themselves in a bank with sound fundamentals. Similarly, the third line represents the private consumption of investors in banks with weak fundamentals, which will receive a bailout \(b_k\) chosen in accordance with (7), and the final line measures the utility from public consumption of all investors in the bad aggregate state.

Using the envelope conditions associated with equations (10) and (13), we can write the first-order condition for this problem in terms of the marginal utilities of consumption of investors in each type of bank,

\[
(1 - q) \left( v'(\tau) - \mu(0; \hat{\pi}) \right) + q \left( v'(\tau - nb) - (1 - n) \mu(0; \hat{\pi}) - n \mu(\bar{\sigma}; \hat{\pi}) \right).
\]

Using equation (7) to replace \(v'(\tau - nb)\) with \(\mu(\bar{\sigma}; \hat{\pi})\), we can re-write this condition as

\[
v'(\tau) = \frac{1 - qn}{1 - q} \mu(0; \hat{\pi}) - \frac{q(1 - n)}{1 - q} \mu(\bar{\sigma}; \hat{\pi}).
\]

If the probability \(q\) of the bad aggregate state were zero, this condition would simply say that the marginal utility of public consumption, \(v'(\tau)\), should be set equal to the marginal utility of private consumption for investors in all banks, \(\mu(0; \pi)\). In this case, the equation reflects a standard Samuelson condition for the efficient provision of the public good. When the probability of the bad state is positive, however, the planner recognizes that there is an additional benefit to raising revenue in period 0: these funds can be used to provide bailout payments in the bad state. The planner will, therefore, set the tax rate higher when \(q\) is positive, and the magnitude of the increase will depend on the marginal utility of private consumption for investors in banks with weak fundamentals, \(\mu(\bar{\sigma}; \hat{\pi})\).

### 3.6 Properties of the constrained efficient allocation

This constrained efficient allocation will serve as an important benchmark in the analysis of decentralized equilibrium in the following sections. Several properties of this allocation are worth emphasizing. The discussion above shows that this allocation is summarized by six numbers, which we denote

\[
\left( \tau^*, \left\{ c^*_1(\sigma), c^*_2(\sigma) \right\}_{\sigma \in \{0, \bar{\sigma}\}}, b^* \right).
\]

The tax rate \(\tau^*\) is chosen according to equation (16) and equates the expected marginal value of public consumption to the expected marginal value of private consumption, taking into account the planner’s (limited) ability to fund private consumption for investors in weak banks through bailouts. The payments \(\left( c^*_1(\sigma), c^*_2(\sigma) \right)\) represent the consumption levels of all im-
patient investors and all patient investors, respectively, in a bank whose idiosyncratic state is $\sigma$, and $b^*$ is the bailout payment made to each bank with weak fundamentals. The following proposition establishes some properties of this allocation.

**Proposition 1.** The constrained efficient allocation satisfies

$$\left(c_1^*(0), c_2^*(0)\right) \gg \left(c_1^*(\bar{\sigma}), c_2^*(\bar{\sigma})\right) \quad \text{and} \quad b^* > 0.$$ 

This result shows that the constrained efficient allocation involves a combination of bailouts and bail-ins at banks that have experienced losses. The bailout $b^*$ gives investors partial insurance against the risk associated with these losses, but the consumption of investors in weak banks remains below that of investors in sound banks; this difference can be interpreted as the degree to which the planner “bails in” the investors in weak banks. The efficient level of insurance is only partial in this environment because offering insurance is costly; it requires the planner to collect more tax revenue, which leads to an inefficiently high level of the public good in the good aggregate state.

One can interpret this allocation as being represented by the following contractual arrangement. Banks offer investors a deposit contract with a face value of $c_1^*(0)$ in period 1 and of $c_2^*(0)$ in period 2. Investors are allowed to withdraw at face value in either period as long as their bank has sound fundamentals and does not experience unusually heavy withdrawals. This contract also has a “bail in” clause: if the bank experiences losses on its asset portfolio, the bank resets the value of its liabilities to investors to $c_1^*(\bar{\sigma})$ in period 1 or $c_2^*(\bar{\sigma})$ in period 2. If early withdrawals $\rho_k$ are expected to exceed $\pi$, the bank also bails in investors as derived in Section 3.3. The analysis above shows that the planner would activate this bail-in clause immediately after the bank’s idiosyncratic state is realized, before any of the period-1 withdrawals take place. The planner would never choose to delay this action by allowing some investors to withdraw $c_1^*(0)$ from a bank that has weak fundamentals or is experiencing a run, because doing so would result in an inefficient allocation of the available resources.

In the next section, we study the equilibrium in the decentralized economy where each bank is operated in the interest of its own investors only. Our focus is on whether private banks will also choose to immediately adjust their liabilities in response to losses on their assets or whether they may in some circumstances choose to delay this action, as well as on how such delay affects the allocation of resources.

## 4 The decentralized economy

In this section we begin our study the decentralized economy. Specifically, in period 0, the investors in each location $k$ form a coalition by pooling their after tax endowments in order to set up “a bank”. In period 0, the payment plan in each bank will be chosen to maximize
the sum of the expected utilities of the investors in the bank. We assume that the investors in all locations can write a complete, state-contingent payment plan (a contract), which could potentially incorporate every scenario that can be faced by their bank.

Compared to the constrained efficient allocation from the previous section, the decentralized economy is different in the following crucial ways. First, banks’ payments are no longer dictated by a fictitious planner who knows everything about their idiosyncratic states, but are chosen by the bank’s themselves. Second, unlike the planner, who cares about economy-wide outcomes, each bank is concern solely with its own investors and views economy-wide variables as being determined exogenously. Third, there is asymmetric information between the banks and the government (who knows the aggregate state, but must wait for \( \theta \) withdrawals to take place before observing bank-specific states). After \( \theta \) withdrawals, each bank will has its own bank-specific state which is a combination of the fundamental state of its assets, the payments it was making during the first \( \theta \) withdrawals and the fraction of the investors in the bank that are impatient and therefore would always choose to withdraw in period 1. At this point in time, the government allocates the existing tax revenues between bailout transfers to banks and the public good. Banks that were bailed-out will have their payment plan for the remaining investors mandated by the government (i.e. these banks are in resolution). The bailout and resolution policy of the government, however, cannot be set up ex-ante, but instead will be chosen to implement the efficient allocation of resources given the available conditions at the time of the intervention.

We show that the constrained efficient allocation cannot, in fact, be sustained in a competitive banking system. A key feature of the equilibrium in the decentralized environment is banks’ delayed response to the crisis: In particular, each bank with weak fundamentals has a private incentive to keep payments high in order to receive a larger bailout transfer later from the government. We show that if payments were adjusted as soon as shocks were realized, then runs do not occur as part of equilibrium. However, we also show that a striking result could obtain in the present setting: the incentive to keep high payment before being bailed out might be so strong that banks choose to tolerate a (partial) run from their investors, instead of adjusting payments at the first sign of trouble.

### 4.1 Contracts with a bail-in option

Denote with \( \bar{c} \) the upper bound on the payment the bank is allowed to set in period 1. We assume that this upper bound \( \bar{c} \) is set by the government. However, unlike in section 3, where the planner can select the exact payment for each investor in all banks, here the government can only set the maximum payment that the bank is allowed to make. Observe that in order to make it feasible for banks to implement the constrained efficient outcome from section 3 the government must set \( \bar{c} \geq c^*_1(0) \). Henceforth, we set \( \bar{c} = c^*_1(0) \). In period 0, each bank selects a state-contingent payment plan, which specifies the payment to investors as they show up to
withdraw as a function of the information available to the bank at the time of the withdrawal:

\[ c^k_1(l, \sigma_k; \rho_k) \in [0, \bar{c}] \]

where \( l \in [0, 1] \) is the fraction of the investors in the bank who had already withdrawn, \( \sigma_k \in \{0, \bar{\sigma}\} \) the fundamental state of the bank’s assets is and \( \rho_k \) is the anticipated fraction of investors that will contact the bank to withdraw in period 1.\(^\text{10}\) In particular, the payment plan can be made fully contingent on the state of the bank, which includes the fundamental state of its assets \( \sigma_k \in \{0, \bar{\sigma}\} \) in addition to any information that can be useful in forecasting the fraction of investors that would choose to withdraw in period 1, \( \rho_k \). Notice that, from the start of period 1, all banks would not only observe the idiosyncratic state of their assets but will also correctly forecast the fraction of investors that will withdraw in period 1 and therefore each bank is perfectly able to make all period 1 payments contingent on the (correct) fraction of investors who would withdraw in the current period.\(^\text{11}\)

Observe that, the payment plan set by the bank can potentially incorporate a *bail-in option*, which allows the bank to lower payment below *face value* \( c^*_1(0) \) whenever this is deemed necessary. Note that the face value is set to the period 1 payment made by sound banks in the constrained efficient allocation. Given that there are no agency cost between the bank and its investors, this bail-in option will be used by the bank whenever this allows it to obtain higher utilities for its investors.\(^\text{12}\)

We can considerably simplified by observing that while the first \( \theta \) withdrawals are being made, no new information becomes available to the bank. We can, therefore, without any loss of generality assume that bank \( k \) makes a common payment \( c^k_1(\sigma_k; \rho_k) \) to each of these investors.\(^\text{13}\) After \( \theta \) withdrawals have occurred and any bailout payments have been made, banks may choose to alter the payments they make to withdrawing investors. We use \( \hat{c}^k_1(\sigma_k; \rho_k) \) to denote the amount given for additional withdrawal in period 1 and \( \hat{c}^k_2(\sigma_k; \rho_k) \) to denote all payments made at \( t = 2 \).

\(^{10}\)In order to economize on notation, we will not explicitly denote the dependence of the allocation on the tax rate \( \tau \) until the end of section 4.5. The choice of the optimal tax rate (which must be made in period 0 before the aggregate state is realized) is postponed until section 4.6.

\(^{11}\)The assumption whereby each bank always holds correct beliefs about the fraction of investors that must be serviced in each period and moreover, is free to adjust its payment whenever these beliefs are revised differentiates our model from the usual approach in the Diamond and Dybvig literature where it is either maintained that banks lack sufficient information to react fast or contracts cannot be (immediately) adjusted in response to changes updates in the beliefs.

\(^{12}\)For example, in period 1, the bank and its investors sign a contract which gives the bank the option to initiate a bail-in clause at any point in period 1. As another interpretation, we can imagine that the bank as having the opportunity to impose a fee on withdrawing investors. In this case, the bank will be able to lower payments to the first \( \theta \) investors from \( c^*_1(0) \) to \( c^*_1(\sigma) \) by imposing a fee of \( c^*_1(0) - c^*_1(\sigma) \) for the duration of the first \( \theta \) withdrawals.

\(^{13}\)See Ennis and Keister (2010a) on this point.
4.2 Banks’ problem

Each bank must choose its payment plan in period 0 - before knowing its idiosyncratic state. Denote with \( p(\bar{\sigma}) \) the probability that bank a given will have weak fundamentals. Conditional on the aggregate state being bad (which occurs with probability \( q \)), each bank will have weak fundamentals with probability \( n \) and therefore, we have \( p(\bar{\sigma}) = qn \). Bank \( k \) chooses its payment plan to maximize the following expression:

\[
\sum_{\sigma_k \in \{0, \bar{\sigma} \}} p(\sigma_k) \left\{ \theta u(c^k_1(\sigma_k; \rho_k)) + (\rho_k(\sigma_k) - \theta)u\left(\hat{c}^k_1(\sigma_k; \rho_k)\right) + (1 - \rho_k(\sigma_k))u\left(\hat{c}^k_2(\sigma_k; \rho_k)\right) \right\}
\]

(18)

The payment plan chosen by each bank must satisfy the following set of conditions:

\[
0 \leq c^k_1(\sigma_k; \rho_k) \leq c^*_1(0) \quad \text{for} \quad \sigma_k \in \{0, \bar{\sigma} \}
\]

(19)

\[
\theta c^k_1(\sigma_k; \rho_k) + (\rho_k(\sigma_k) - \theta)\hat{c}^k_1(\sigma_k; \rho_k) + (1 - \rho_k(\sigma_k))\frac{\hat{c}^k_2(\sigma_k; \rho_k)}{R} \leq (1 - \tau) + b_k
\]

(20)

\[
u'\left(\hat{c}^k_1(\bar{\sigma}; \rho_k)\right) = R u'\left(\hat{c}^k_2(\bar{\sigma}; \rho_k)\right) \leq v'\left(\tau - \int_0^1 b_k(\sigma_k, c_1)\right) \quad \text{with} \quad " = " \quad \text{if} \quad b_k > 0
\]

(21)

\[
b_k \geq 0 \quad \text{iff} \quad k \in L
\]

(22)

According to (20), the payment plan must satisfy the bank’s budget constraint for each realization of its idiosyncratic state \( \sigma_k \in \{0, \bar{\sigma} \} \) and given any fraction of investors choosing to withdraw in period 1. Conditions (21) and (22) represent the constraints on the bank imposed by the bailout and resolution policy of the government. First, the payment plan for the remaining investors in banks that were bailed out will be selected by the government in order to satisfy the first order condition in (21) and according to (22) only banks with weak fundamentals could potentially receive a positive bailout transfer from the government. Next, we proceed to bank’s payment plan under two circumstances. First, the payment plan mandated by the government conditional on the bank being placed in resolution after the first \( \theta \) withdrawals. Second, the payment plan conditional on the bank having sound fundamentals.

Resolution. Each bank that is bailed out is placed in resolution and its payment plan from now on is selected by the government. We have the following result:

**Proposition 2.** In any decentralized equilibrium,

(i) A bank would not experience run after being placed in resolution.
(ii) All banks in resolution set the same post-bailout payment plan determined by (21).

Observe from (21) that all banks with a bailout will have their post-bailout payment set to a common level \( \hat{c}_1(\sigma), \hat{c}_2(\sigma) \) regardless of the payments \( c_1^k(\sigma_k; \rho_k) \) that they were making before being bailed out (as long as these payments lead to a bailout). This common post-bailout level of payments will be characterized by the following condition:

\[
u' (\hat{c}_1(\sigma)) = Ru' (\hat{c}_2(\sigma)) = v' \left( \tau - \int_0^1 b_k \right) \quad (23)
\]

The bailout policy of the government therefore implies that after being bailed out, the payment plan to the remaining investors is determined solely by economy-wide conditions which are beyond the control of any given bank. In particular, each bank views \( \hat{c}_1(\sigma), \hat{c}_2(\sigma) \) as an exogenously specified component of its payment plan which will be activated as soon as the bank receives a bailout from the government. Moreover, from (23) we have \( \hat{c}_1(\sigma) < \hat{c}_2(\sigma) \) i.e once a bank is in resolution, patient investors would strictly prefer to withdraw in period 2 and therefore withdrawals that still occur in period 1 will be made solely by the impatient investors. The results so far in this section are summarized in the next proposition.

**Sound fundamentals.**Conditional on having sound fundamentals \( \sigma_k = 0 \), the bank’s payment plan would be set to satisfy the first order condition in (24) and the resource constraint in (20):

\[
u' \left( c_1^k(0; \rho_k) \right) = u' \left( \hat{c}_1^k(0; \rho_k) \right) = Ru' \left( \hat{c}_2^k(0; \rho_k) \right) \quad (24)
\]

Note that for any given withdrawal demand function \( \rho_k \), the payment plan will satisfy the condition in (24), which implies that waiting until period 2 to withdraw is strictly preferred by patient investors in case their bank’s fundamentals are sound:

\[c_1^k(0; \rho_k) = \hat{c}_1^k(0; \rho_k) < c_2^k(0; \rho_k)\]

therefore if \( \rho_k \) is to be consistent with equilibrium it must specify that only impatient investors withdraw in period 1 in case the bank has sound i.e. fundamentals \( \rho_k(\sigma_k) = \pi \) for \( \sigma_k = 0 \).

**Proposition 3.** Conditional on their fundamentals being sound, banks do not experience a run and select a payment which, for any given \( \tau \), is identical to the one selected by the planner in section 3.

According to Proposition 3, all banks with sound fundamentals will behave identically and moreover, their payment plan exactly coincides with the payment that would have been selected had they been operated by the planner from section 3. Propositions 2 and 3 considerably simplify the analysis by showing that all of the interesting activity in the decentralized economy occurs while banks with weak fundamentals are making payments to the first \( \theta \) fraction
of their investors. In order to fully characterize the bank’s problem, we still need to specify investors’ withdrawal behavior conditional on the payment plan that was selected in period 0.

4.3 Withdrawal strategies

For any given payment plan chosen by the bank in period 0, investors will play a withdrawal game in period 1 and 2, the outcome of which is correctly predicted by the bank at the time the payment plan was being chosen. Investor $i$ in bank $k$ chooses a strategy of the form

$$y^k_i : \Omega \times \Sigma \rightarrow \{0, 1\},$$

which assigns a withdrawal decision to each combination of her preference type $\omega^k_i \in \{0, 1\} = \Omega$ and her bank’s idiosyncratic state $\sigma^k \in \{0, \bar{\sigma}\} = \Sigma$. Here $y^k_i = 0$ corresponds to withdrawing early and $y^k_i = 1$ corresponds to waiting to withdraw in the final period. Let $y^k$ denote the profile of withdrawal strategies for all investors in bank $k$. Using this profile, we can derive the function:

$$\rho_k : \Sigma \rightarrow [0, 1],$$

which gives the measure of early withdrawals bank $k$ will face in each state when its investors follow the strategy profile $y^k$. The no-run (or truth telling) strategy for each investor in bank $k$ is the following:

$$y^k_i \left(\omega^k_i, \sigma^k\right) = \omega^k_i$$

(25)

That is, an investor withdraws in period 1 only if impatient, in which case we have $\rho_k(\sigma^k) = \pi$, i.e. the fraction of investors withdrawing in period 1 equals the fraction of impatient depositors in the bank. Next, in order to formulate a profile of withdrawal strategies which generates a run, we will make use of the fact that any payment plan consistent with equilibrium in the decentralized economy must satisfy a number of properties. First, when impatient, any given investor has a strictly dominant strategy to withdraw in period 1, regardless of the bank’s payment plan. That is, since withdrawing early is a dominant action for impatient investors, $\rho_k(\sigma^k) \geq \pi$ will hold for every $\sigma^k$. Second, from (23) we know that after their bank has been bailed out, patient investors strictly prefer to withdraw in period 2. Third, from Proposition 3, all banks with sound fundamentals set a payment plan identical to section 3 and investors in these banks withdraw in period 1 only if impatient. By combining all of the above, we can formulate the following partial run strategy

$$y^k_i \left(\omega^k_i, \sigma^k\right) = \begin{cases} 
\omega^k_i & \text{if } \sigma^k = 0 \\
\omega^k_i & \text{if } \sigma^k = \bar{\sigma} \text{ and } i > \theta \\
0 & \text{if } \sigma^k = \bar{\sigma} \text{ and } i \leq \theta
\end{cases}$$

(26)
That is, if a bank run is to be consistent with equilibrium, it must be restricted to investors with an opportunity to be among the first $\theta$ to withdraw in banks with weak fundamentals.

**Proposition 4.** All investors in any bank that could experience a run in equilibrium must follow the partial run strategy in (26).

So, suppose that bank $k$ turns out to have weak fundamentals in period 1 and consider an investor with an opportunity to be among the first $\theta$ to withdraw. When patient, she is best responding with the strategy in (26) whenever the following condition is satisfied:

$$c_1^k(\bar{\sigma}; \rho_k) > \hat{c}_2^k(\bar{\sigma}; \rho_k)$$  \hspace{1cm} (27)

Therefore, if the bank selects a payment plan such that (27) is not satisfied, then each investor in the bank would follow the no-run strategy in (25). In contracts, if the selected payment plan satisfies (27), then all investors in the bank would follow the partial run strategy in (26). Therefore, for any given payment plan selected by the bank in period 0, consistency requires that the fraction of investors withdrawing in period 1 is the following:

$$\rho_k(\sigma_k) = \begin{cases} 
\pi & \text{if } \sigma_k = \bar{\sigma} \text{ and } c_1^k(\bar{\sigma}; \rho_k) \leq \hat{c}_2^k(\bar{\sigma}; \rho_k) \\
\pi & \text{and } c_1^k(\bar{\sigma}; \rho_k) > \hat{c}_2^k(\bar{\sigma}; \rho_k) \\
\theta + (1 - \theta)\pi & \text{if } \sigma_k = \bar{\sigma} \text{ and } c_1^k(\bar{\sigma}; \rho_k) \leq \hat{c}_2^k(\bar{\sigma}; \rho_k) \\
\theta + (1 - \theta)\pi & \text{and } c_1^k(\bar{\sigma}; \rho_k) > \hat{c}_2^k(\bar{\sigma}; \rho_k) 
\end{cases}$$  \hspace{1cm} (28)

Finally, at the time payments plans are selected in period 0, the bank would use (28) in order to anticipate the fraction of its investors choosing to withdraw in period 1 and 2 respectively. In other words, the function $\rho_k$ will be treated as an additional constraint on the set of feasible payment plans available to the bank. Each bank will therefore choose its payment plan in order to maximize the sum of the expected utilities of its investors (18) subject to (19), (20), (21), (22) and (28).

Before proceeding to characterize the equilibrium outcome, note that the bank’s problem can be restated in the following equivalent way. In period 0 the bank selects a collection of payment schedules for any given combination of the idiosyncratic state of its assets $\sigma_k \in \{0, \bar{\sigma}\}$ and the fraction of investors withdrawing in period 1, $\rho_k \in [\pi, 1]$. In addition, any given payment schedule selected by the bank must satisfy the conditions in (19), (20), (21) and (22). A collection of payment schedules – one for each combination of $(\sigma_k, \rho_k)$ will be called a payment plan. In period 1, the bank and all of its investors observe the realization of $\sigma_k$. In addition, each investor observes whether she is patient or impatient and her order in the opportunities to withdraw and then chooses to withdrawing either in period 1 or period 2, which would result in period 1 withdrawal demand equal to $\rho_k^*(\sigma_k) \in [\pi, 1]$. This is is where our assumption whereby the bank can always correctly predict the actual fraction of investors choosing to withdraw in period 1 becomes important, since this will allow it use the “correct”
payment schedule (i.e. the one associated with \((\sigma_k, \rho^*_k(\sigma_k))\)) in order to service all investors from the start of period 1. Finally, for any given payment schedule, investors must be best responding with their withdrawal choices. It is not hard to see that this formulation of the problem is equivalent to the one given above.

4.4 Choice of early payments

In this section, we characterize the bank’s choice of payment during the first \(\theta\) withdrawals. First, any bank with sound fundamentals pays \(c^*_1(0)\) to all investors withdrawing in period 1 and \(c^*_2(0)\) to all investors withdrawing in period 2, where these payments are characterized by (20) and (24). Since \(c^*_1(0) < c^*_2(0)\), it follows that conditional on having sound fundamentals, only the impatient investors will show up to withdraw in period 1 and the expected utility for the investors in the bank in this case will be:

\[
V (c^*_1(0); 0) = \pi u (c^*_1(0)) + (1 - \pi) u (c^*_2(0))
\]

Next, suppose that the equilibrium involves bailouts for banks with weak fundamentals. That is, after \(\theta\) withdrawals, the government provides a bailout of \(b > 0\) to all banks with weak fundamentals and hereafter sets their payment plan to \((\hat{c}_1(\bar{\sigma}), \hat{c}_2(\bar{\sigma}))\) to satisfy the condition for the ex-post optimal bailout policy in (23). Finally, in order to complete the description of the decentralized equilibrium, we must characterize the choice of early payments conditional on the bank having weak fundamentals, so let \(c^*_{DE}(\bar{\sigma}) \in [0, c^*_1(0)]\) denote the payment during the first \(\theta\) withdrawals made by banks with weak fundamentals.

**Optimal choice of \(c^*_{DE}(\bar{\sigma})\).** Next, we focus on a particular choice of early payments: \(c^*_{DE}(\bar{\sigma}) = c^*_1(0)\) and derive conditions which ensure this is indeed consistent with equilibrium. In order to show that banks with weak fundamentals choose to set \(c^*_1(0)\) we must compare this strategy for setting early payments with two alternatives plans of action, to which we turn next. First, instead of setting \(c^*_1(0)\) when its fundamentals are weak, the bank in question can choose the following payment plan: set \(\hat{c}_2(\bar{\sigma})\) to the first \(\pi - \theta\) impatient and \(\hat{c}_2(\bar{\sigma})\) to the \(1 - \pi\) patient investors. Observe that, if \(c^*_1(0) > \hat{c}_2(\bar{\sigma})\) then this payment plan would have the desirable property of preventing a run at the cost of, however, of having to give lower payment to the first \(\theta\) investors. Conditional on the bank having weak fundamentals, the ex-ante utility associated with this alternative strategy is:

\[
V (\hat{c}_2(\bar{\sigma}); \bar{\sigma}) \equiv \theta u (\hat{c}_2(\bar{\sigma})) + (\pi - \theta) u (\hat{c}_1(\bar{\sigma})) + (1 - \pi) u (\hat{c}_2(\bar{\sigma}))
\]

(29)

There is also a third strategy available to the bank which is to entirely avoid a bailout. In particular, let \(\hat{c}_1\) denote the cut-off value of early payments, such that for any payment below this cut-off the government chooses not to bail out the bank (that is, (21) holds with strict inequal-
ity whenever \( c_i^k < \tilde{c}_1 \). Next, let \( c_i^{NB}(\tilde{\sigma}) \) denote the optimal choice of early payments among those which are less or equal to \( \tilde{c}_1 \). Hence, conditional on having weak fundamentals, the ex-ante utility associated with the optimal no-bailout plan will be determined by the following program:\(^{14}\)

\[
V \left( c_1^{NB}(\tilde{\sigma}); \tilde{\sigma} \right) \equiv \begin{cases} 
\max_{\{c_1, \hat{c}_1, \hat{c}_2\}} & \theta u(c_1) + (\pi - \theta) u(\hat{c}_1) + (1 - \pi) u(\hat{c}_2) \\
\text{s.t.} & \theta c_1 + (\pi - \theta) \tilde{c}_1 + (1 - \pi) \frac{\hat{c}_2}{R} \leq (1 - \tau)(1 - \tilde{\sigma}) \\
& u'(\hat{c}_1) = Ru'(\hat{c}_2) \leq v'(\tau - nb)
\end{cases}
\] (30)

Finally, banks with weak fundamentals best respond by choosing \( c_i^*(0) \) whenever the following weak inequality is satisfied:

\[
V \left( c_i^*(0); \tilde{\sigma} \right) \geq \max \left\{ V \left( \hat{c}_2(\tilde{\sigma}); \tilde{\sigma} \right), V \left( c_{i}^{NB}(\tilde{\sigma}); \tilde{\sigma} \right) \right\}
\] (31)

The above condition represents a necessary condition for the allocation where all banks set \( c_i^*(0) \) during the first \( \theta \) withdrawals (regardless of their state \( \sigma_k \)) to be consistent with equilibrium. If, on the other hand, (31) is not satisfied, then in equilibrium banks would choose to follow one of the two alternative options by either setting \( \hat{c}_2(\tilde{\sigma}) \) or \( c_i^{NB}(\tilde{\sigma}) \) to the first fraction of \( \theta \) investors.

### 4.5 Equilibrium for given \( \tau \)

We are now ready to fully characterize the properties of the decentralized equilibrium, but before doing so, we re-introduce the dependence of the allocation on the choice of \( \tau \) made by the government in period 0. That is, we will write \( c_i^*(0, \tau) \) and similarly for the remaining variables. We begin by asking whether the constrained efficient outcome from section 3 can be implemented in the decentralized setting of the current section. In particular, a bank with sound fundamentals must pay \( c_i^*(0, \tau^*) \) in period 1 and \( c_2^*(0, \tau^*) \) in period 2, whereas a bank with weak fundamentals must pay \( c_i^*(\tilde{\sigma}, \tau^*) \) in period 1 and \( c_2^*(\tilde{\sigma}, \tau^*) \) in period 2 (while also receiving a bailout of \( b^* \) from the government). The answer turns out to be negative:

**Proposition 5.** The constrained efficient allocation cannot be sustained in a decentralized economy.

Once we have set up the environment in the previous sections, the proof of this result becomes relatively straightforward. If we want to implement the constrained efficient outcome,
all banks must be setting the same state-contingent payment that was chosen by the planner in section 3. However, the constrained efficient outcome is not an equilibrium here because a marginal increase $\varepsilon > 0$ in the payment to the first $\theta$ investors will not change their withdrawal behavior (as long as $c_1^*(\bar{\sigma}, \tau^*) + \varepsilon < c_2^*(\bar{\sigma}, \tau^*)$ and, at the same time, will not penalize the remaining $1 - \theta$ investors in the bank since their payment plan would remained unchanged and equal to $\left(c_1^*(\bar{\sigma}, \tau^*), c_2^*(\bar{\sigma}, \tau^*)\right)$. In other words, we have shown that a given bank $k$ has a profitable deviation from the payment plan required to implement the constrained efficient outcome and by doing so we have established the desired result in above proposition. Given that the constrained efficient allocation cannot be sustained in equilibrium, we must look for a different outcome. Proposition 6 shows that for any given value of the tax rate $\tau$, the equilibrium in the decentralized environment will be one of three types, with the crucial difference being in the early choices made by banks with weak fundamentals.

**Proposition 6.** In any equilibrium of the decentralized economy we have:

(i) $c_{DE}^1(\bar{\sigma}, \tau) \in \{c_1^*(0, \tau), \hat{c}_2(\bar{\sigma}, \tau), c_1^{NB}(\bar{\sigma}, \tau)\}$

(ii) In equilibrium, investors follow the no-run strategy in (25) when either $c_1^*(0, \tau) \leq \hat{c}_2(\bar{\sigma}, \tau)$

is true, or the optimal choice of early payments is $c_{DE}^1(\bar{\sigma}, \tau) \in \{\hat{c}_2(\bar{\sigma}, \tau), c_1^{NB}(\bar{\sigma}, \tau)\}$.

(iii) In equilibrium, investors follow the partial run strategy in (26) whenever $c_{DE}^1(\bar{\sigma}, \tau) = c_1^*(0, \tau) > \hat{c}_2(\bar{\sigma}, \tau)$.

Notice that in case (iii) conditional on having weak fundamentals, a bank sets $c_1^*(0, \tau)$ and thus chooses not to undertake any payment adjustment whatsoever before being bailed out and placed in resolution by the government. Notice that from a government perspective, the above behavior would render banks with weak fundamentals indistinguishable from banks with sound fundamentals for the duration of the first $\theta$ withdrawals. Next, we establish the existence of parameter values $(R, \pi, \theta, \gamma, \delta, n, q)$ and a tax rate $\tau$ such that banks indeed choose to set $c_{DE}^1(\bar{\sigma}, \tau)$ equal to $c_1^*(0, \tau)$.

**Proposition 7.** There exist parameter values and a tax rate $\tau$ such that all banks (including those with weak fundamentals) are paying $c_1^*(0, \tau)$ during the first $\theta$ withdrawals and all investors follow the partial run strategy in (26).

Proposition 7 is demonstrated on figure 2 where investor preferences are assumed to be CRRA.\textsuperscript{15}

\textsuperscript{15} We use the following parameter values for the example in figure 2: $R = 1.05, \pi = 0.5, \theta = 0.25, \gamma = 5, \delta = 0.1, q = 0.02, n = 0.25, \bar{\sigma} = 0.5$
\[ \frac{(c_1 + \omega c_2)^{1-\gamma}}{1-\gamma} \quad \text{and} \quad \delta \frac{g^{1-\gamma}}{1-\gamma} \]

The solid line on panel (b) plots investors’ incentive to run when their bank sets \( c_1^{DE}(\bar{\sigma}, \tau) = c_1^*(0, \tau) \):

\[ f_I(\tau) \equiv c_1^*(0, \tau) - \hat{c}_2(\bar{\sigma}, \tau) \]

If \( f_I(\tau) > 0 \), then patient investors with an opportunity to be among the first \( \theta \) to withdraw would run whenever their bank has weak fundamentals, anticipating to be paid less if they wait until the last period to withdraw. Next, the solid line on panel (c) plots banks’ incentive to choose \( c_1^*(0, \tau) \) relative to the remaining two options available:

\[ f_B(\tau) \equiv V \left( c_1^*(0, \tau); \bar{\sigma}, \tau \right) - \max \left\{ V \left( \hat{c}_2(\bar{\sigma}, \tau); \bar{\sigma}, \tau \right), V \left( c_1^{NB}(\bar{\sigma}, \tau); \bar{\sigma}, \tau \right) \right\} \]

If \( f_B(\tau) > 0 \), then banks with weak fundamentals will have a dominant strategy to choose \( c_1^*(0, \tau) \). Note that if there exist a range of \( \tau \) where both \( f_I(\tau) \) and \( f_B(\tau) \) are greater than zero (as is the case in Figure 2), it then follows that for each \( \tau \) belonging to these range we can construct an equilibrium involving both a delayed response on the part of the banks and a run on the part of the investors.

### 4.6 The choice of \( \tau \)

It still remains to characterize the equilibrium of the overall game, which involves the government choosing the tax rate in period 0 in period 0, before knowing the realization of the aggregate state \( s \in \{G, B\} \). The optimal choice of the tax rate, denoted \( \tau^O \), is depicted by the vertical dashed line and it is associated with the sub-game which maximizes the sum of ex-ante expected utilities of the investors. We can show the following:

**Proposition 8.** There exist parameter values in the overall game such that all banks (including those with weak fundamentals) are paying \( c_1^*(0, \tau) \) during the first \( \theta \) withdrawals and all investors follow the partial run strategy in (26).

In particular, for any given choice of \( \tau \), the government anticipates both the payment plan chosen by the banks and the withdrawal strategies chosen by the investors and then selects \( \tau \) in order to maximize the utility of an representative investor \( W(\tau) \), where:

\[ W(\tau) \equiv \frac{(1-q) \left\{ V \left( c_1^*(0, \tau); \bar{\sigma}, \tau \right) + v(\tau) \right\}}{+q \left\{ (1-n)V \left( c_1^{DE}(\bar{\sigma}, \tau); \bar{\sigma}, \tau \right) + nV \left( c_1^*(0, \tau); \bar{\sigma}, \tau \right) + v(\tau-nb(\tau)) \right\}} \]

That is, with probability \( 1-q \) the aggregate state in period 1 will be \( G \), in which case all banks have sound fundamentals and no-bailouts will be made. On the other hand, with probability \( q \),
the aggregate state will be $B$ in which case a fraction $n$ of the banks will have weak fundamentals. Conditional on having weak fundamentals, each bank would choose a payment for the first $\theta$ investors $c_{1}^{DE}(\bar{\sigma}, \tau)$ and then, at the time all bank-specific information becomes known to the government, would receive a bailout whenever the shadow value of the resources available to service the remaining $1-\theta$ investors is below the marginal utility of the public good. We can see from panel (a) in the figure, that $f_{I}(\tau^{D}) > 0$ and $f_{B}(\tau) > 0$ and therefore the equilibrium in the overall game involves both a strategic delay and a run as stated by Proposition 8.

![Figure 2:](image)

### 4.7 Properties of the decentralized equilibrium

Here, we provide a discussion of the properties of the decentralized economy and especially relate the features of the decentralized outcome to the features exhibited by the constrained efficient outcome. In addition, we show that the primary results are robust to a number of variations in the underlying environment.

**Bailouts delay bail-ins.** Note that the constrained efficient outcome failed to be an equilibrium here because a bank with weak fundamentals would never find it optimal to lower payments to the first $\theta$ investors all the way down to $c_{1}^{*}(\bar{\sigma}, \tau^{*})$ as prescribed by the constrained efficient payment plan (i.e. a complete bail-in is never individually optimal). Henceforth, we refer to such behavior as a strategically delayed response on the part of the banks. Indeed, we will show that the equilibrium in this environment generally involves strategically delayed responses. If banks engage in this sort of strategic delay then they fully anticipate that their post-bailout payment in period 1 would actually be adjusted downwards compared to their current level. Note also that the adjective “strategic” here is crucial since it signifies that it is the bank’s own choice to delay their reaction to bad news and keep current payments at their unsustainably high level. That is, the bank could have chosen and implemented a payment plan such that payments to all investors are immediately and fully adjusted in response to any
losses. Also, notice that the government is assumed to lack commitment which implies that the bailout policy will be characterized by the single condition in (21). In other words, it will not be ex-post optimal to punish banks giving large payments (by, for example refusing to bail them out or imposing a penalty for such behavior) and therefore any treat from the government to do otherwise fails to be credible.

**The banks can always avoid a run, but may decide not to.** Throughout the entire analysis, we have always maintained the assumption that banks always have sufficient information about the state of the world and are able to write a fully state-contingent contract which would take into account each possible future scenario. These two conditions alone ensure that, in principle, a run-proof contract can always be written and subsequently implemented. In other words, each bank has, on its disposal, all of the tools necessary to prevent runs. Banks and their investors, however, could deliberately choose not to write such a contract in order maximize their benefit from the government’s bailout intervention. The equilibrium, therefore, will be characterized by rigid contractual arrangement and banks behaving as if they lacked sufficient information to respond to an unfolding crisis. This outcome, moreover, arises fully endogenously.

Indeed, observe from Propositions 7 and 8 that the bank anticipates that by keeping payments at \( c_1^*(0, \tau) \) a bailout and a subsequent resolution are going to be put in place after the first \( \theta \) withdrawals. In this case, the only way the bank can derive a benefit by lowering payments before the bailout (i.e. instituting an immediate bail-in) is when such an action would affect investors’ withdrawal behavior. In particular, suppose \( c_1^*(0, \tau) > \hat{c}_2(\bar{\sigma}, \tau) \), in which case, the bank is fully aware that continuing to pay \( c_1^*(0, \tau) \) while the pre-bailout withdrawals are taking place would generate a run. Indeed, a patient investors, knows that after being bailed out and placed in resolution, her bank would reschedule payments right away and start paying \( \hat{c}_1(\bar{\sigma}, \tau) \) to those withdrawing in period 1 and \( \hat{c}_2(\bar{\sigma}, \tau) \) to those withdrawing in period 2. In this case, if \( c_1^*(0, \tau) > \hat{c}_2(\bar{\sigma}, \tau) \), any patient investor with an opportunity to be among the first \( \theta \) to withdraw knows that she can withdraw before the bank is placed in resolution and receive \( c_1^*(0, \tau) \) which is strictly better compared to waiting until the last period in order to be paid \( \hat{c}_2(\bar{\sigma}, \tau) \) and therefore, a run would ensue. Indeed, the run on the bank can be viewed as the cost that the bank pays for delaying its payment adjustment. In other words, each bank would always prefer to avoid a run. However, under certain conditions, taking actions to avoid a run (i.e. bailing-in investors who currently withdraw) might be even less desirable from the perspective of banks that had suffered losses.

**Commitment on the part of the banks.** The equilibrium in section 4 was derived assuming that banks can pre-commit to their payment plans in period 0. As a result, banks choices in period 0 can be used to influence investors’ withdrawal decisions in the periods to follow. In contrast, if banks were unable to commit to their payment plans ex-ante, then the prob-
lem could be formulated as follows: the investors in each bank first choose their withdrawal strategies and then banks would best respond by choosing a payment plan given this profile of withdrawal strategies. Observe that the substantial difference in this case is that banks would choose their payment plan taking as given investors’ withdrawal choices (instead of potentially trying to influence their choices as in section 4.3).

Notice, however, that removing commitment on the part of the banks would, in fact, leave the most striking implication of the model unchanged. In particular, suppose that in the equilibrium of the game with commitment all banks set \( c_1^* (0, \tau) \) to the first \( \theta \) withdrawals and all investors follow the partial run profile in (26). Therefore, it must be the case that, conditional on having weak fundamentals, each bank prefers to set \( c_1^* (0, \tau) \), even when a run can be prevented by decreasing early payments by a sufficient amount.

Now, assume that the bank cannot commit to their payment plan. In this case, investors partial run strategy in (26) would be taken as given and therefore each bank with weak fundamentals would have an even stronger incentive to set \( c_1^* (0, \tau) \) (since, investors’ strategies will be treated as fixed by the banks and therefore lowering payments to any level below \( c_1^* (0, \tau) \) would not induce them to switch to the no-run strategy in (25)). The conclusion is that the outcome where all banks pay the same amount to the first \( \theta \) withdrawals and all investors follow the partial run profile in (26) is an equilibrium of the game without commitment whenever it is an equilibrium of the game with commitment (not that the reverse is not necessarily true).

5 Real shocks, sunspots, and panics

We begin this section by showing that financial panic in this environment cannot be explained solely as a self-fulfilling runs. Rather, what is necessary is the occurrence of a fundamental shock which would then induce banks to delay their reaction and thus give an incentive to the investors to withdraw while payments are being kept at an unsustainable level.

**Proposition 9.** Runs cannot occur unless the shock to fundamentals \( \bar{\sigma} \) is sufficiently large.

The intuitive is that the required payment adjustment after sustaining relatively small losses is itself relatively small and therefore the miss-allocation of resources resulting from strategic delay is not enough to induce investors to withdraw before the government’s bailout intervention. Next, we show that, regardless of the size of the shock \( \bar{\sigma} \), runs will still not take place unless the negative shock to fundamentals is also sufficiently widespread.

**Proposition 10.** Runs cannot occur unless bank’s losses are sufficiently widespread (i.e. \( n \) is large enough).

According to Proposition 6, banking panics would not occur unless the anticipated haircuts are severe enough to justify them in the first place. If the fundamental shock ends up being restricted to a small fraction of assets in the banking sector, then the government would
have enough capacity to essentially provide deposit insurance to the affected banks, without imposing drastic cuts in the provision of the public good (see the middle panel on figure 4).

Finally, when $\theta$ is sufficiently small, the government infers the full state of the financial system almost immediately and runs do not occur as part of equilibrium. In fact, runs in this setting could take place only because banks postpone adjusting their payments to reflect their true financial state. The longer banks engage in this sort of strategic delay, the larger is the resulting miss-allocation of resources and hence the bigger is the incentive for investors to withdraw early (i.e. run on the banks). This, in turn, highlights an important feature of the model – the timing of the bailout intervention is of crucial importance. In particular, an optimally timed bailout plan – one taking place just after banks have sustained losses but before they begin to make payments to investors is capable of eliminating runs while preserving the socially valuable contingent transfer purpose of the bailout program.

**Proposition 11.** *If the government is able to learn banks’ idiosyncratic states and implement the bailout program at the start of period 1, before banks have had the opportunity to engage in strategic delay (i.e. $\theta = 0$), then there are no runs and the decentralized allocation is constrained efficient.*

Hence, a desirable arrangement in this setting is one where the government manages to collect information about banks’ losses and implement bailout program before the banks have had an opportunity to engage in fragility-enhancing actions such as strategic delay.

To summarize the results in this section: in this environment a banking panic could not occur unless all of the following conditions are satisfied: (i) the losses in bank’s balance sheets are non-trivial, (ii) these losses are sufficiently wide-spread and (iii) the government’s bailout intervention occurs after banks had made payments to some of their investors while anticipating to be bailed out. Also, note the novel features of our environment, which differentiate it from the existing models of banks runs. A run in this environment is not explained by exogenously imposed rigidity in the contractual arrangements which prevent the bank from fully adjusting its payment plan to the withdrawal demand. Nor is it explained by informational frictions, which prevent the bank from swiftly adjusting payments in response to certain events (like a bank run). Instead, a run could be viewed as the cost banks pay for engaging in strategically delayed responses, whose aim it is to secure a larger bailout later on.

### 5.1 Discussion

From a social perspective a longer delay (in payment adjustment) would make both the taxpayer and the remaining investors in the bank (those that withdraw relatively late) worse off. However, the degree to which the government is capable of resolving the tension between what is ex-ante desirable and ex-post optimal depends (at least in part) on some exogenous constraints elsewhere in the economy (i.e. in the public sector or in the political arena) which
determine to what extent the government is able to ex-post deviate from its ex-ante optimal policy. Thus, the slow reaction of the government can also be interpreted to reflect political power from certain special interest groups. For example, investors who are well-connected politically might also be among the first with an opportunity to withdraw during times of financial distress. These investors could then use their political influence to convince the policy maker not to step in right away and instead to impose the haircuts on those withdrawing later on.

In addition, the timing of the government intervention might reflect opaque incentives faced by regulators. Kroszner and Strahan (1996) show that throughout the eighties the Federal Savings and Loan Insurance Corporation (FSLIC) was faced with a severe shortage of cash with which to resolve insolvent thrift institutions. This lack of funds forced the FSLIC to practice regulatory forbearance and to delay its explicit intervention in insolvent mutual thrifts in anticipation that the government would eventually supply additional resources. This delay led a large number of insolvent thrift institutions to maximize the value of future government liabilities guarantees (at the taxpayers’ expense) by continuing to pay high dividends until the eventual resolution mechanism was put in place.

Another reason for a delay in the banking resolution might be related to political timing. Brown and Dunc (2005) study episodes of government resolutions of failed banking institutions in 21 major emerging markets during the 1990s and provide evidence that the timing of the government’s intervention depends on the electoral cycle. In particular, costly government interventions which would impose a high cost on the taxpayer and would also fully reveal the extent of the financial crisis (and thus may raise questions of how the government allowed this to happen in the first place) were significantly less likely to occur before elections. In other words, political factors might lead inaction or to a delay in the adoption of beneficial economic policies (Rogoff and Sibert 1988).

6 Concluding remarks

The existing literature has generated banks runs by assuming that banks are slow to react to an incipient run. This slow reaction gives investors a strong incentive to withdraw early (run on the bank) anticipating that future payments must be adjusted to reflect the bank’s state of financial distress. Therefore, from the perspective of an individual investor, a bank run can be seen as a preemptive action whose goal is to avoid a future haircut from a bank that is currently being too optimistic with respect to its future prospects. In the theoretical models of bank runs the reason for this slow response has been exogenously imposed rather than endogenously derived as part of the equilibrium. In particular, banks’ slow reaction has been rationalized by either assuming that (i) contracts are rigid and therefore cannot be altered (at least not right away) in the event of a run or (ii) the bank cannot respond efficiently to a run because it lacks
sufficient information. For example, the bank might be unaware that a panic is underway before withdrawals become unusually high and hence it might be too slow in reacting to the run.

Both (i) and (ii) are questionable assumptions and especially so when it comes to systemic banking panics. First, the rigid contracts assumption (also referred to as the “simple contracts” framework) simplifies the model and allows for greater tractability. Moreover, when times are normal, contracts do appear to be fixed e.g. a demand deposit contract when the bank is not experiencing a run. In fact, given that a systemic panic is by definition a rare event, the “rigidness” of the financial arrangements is only rarely put to the test. Nonetheless, when these systemic events do occur, renegotiation of financial commitments appears to be the norm rather than the exception (e.g. Argentina, Island, Cypriot, US, and others). Second, assuming that investors become aware of a run before their bank does is also not particularly appealing. In particular, there must be an event which is observed by a large fraction of the investors in the bank and, but not by the bank itself (at least not right away). This event can then serve as a coordination device for investors’ decision to withdraw early or to wait.

One of the primary objectives of our analysis is to provide a theory of banking panics which does not suffer from these shortcomings. For that purpose, we have presented a model of financial crisis where three are no frictions between the bank and the investors. In particular, (i) banks are benevolent and want to maximize the utilities of their investors. (ii) Banking arrangements are fully flexible and, at any point in time, can be altered in reaction to the prevailing situation and (iii) Information related to the future prospects of the bank would become known to the bank at the same time (or before) it becomes known to the investors. Financial panics occur because banks in financial distress find it optimal to maximize the value of the government’s guarantees by initially keeping payments relatively high in order to hide from the regulator the fact that they have suffered losses. Investors in turn, realizing that the period of high payments will be relatively short lived and entirely due to the moral hazard generated by government’s bailout program rush to withdraw (run on the bank) before the government’s steps and implements measures to resolve the problematic institutions (these measures will involve haircuts to those still invested in the problematic banks).

The model delivers a number of predictions considered to be empirically plausible. Panics based purely on self-fulfilling expectations cannot arise. Rather, what is necessary is a real shock, which in addition must be of sufficient severity and scope, in order to create the necessary conditions for strategic delay on the part of the banks and runs on the part of the investors. On the other hand, a relatively mild shock (or one that is restricted only to a small fraction of the banks) would lead to a situation akin to deposit insurance in normal times, i.e., transfers to the banks affected will be made, but such an intervention will be implemented in an orderly manner without delays and panics. Finally, banking panics do not occur unless the government is expected to intervene and provide bailouts.
Appendix A

To be added ...

References


