Discussion Paper Series – CRC TR 224

Discussion Paper No. 091
Project C 03

Bailouts, Bail-ins and Banking Crises

Todd Keister*
Yuliyan Mitkov**

May 2019

*Rutgers University, todd.keister@rutgers.edu
**Adenauerallee 24-26, 53113 Bonn. +49 (0)228 73 79 97. Email: ymitkov@uni-bonn.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Bailouts, Bail-ins and Banking Crises*

Todd Keister
Rutgers University
todd.keister@rutgers.edu

Yuliyan Mitkov
University of Bonn
ymitkov@uni-bonn.de

October 15, 2018

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. Banks in our model hold risky assets and are able to write complete, state-contingent contracts with investors. In the constrained efficient allocation, banks experiencing a loss immediately bail in their investors and this bail-in removes any incentive for investors to run on the bank. In a competitive equilibrium, however, banks may not enact a bail-in if they anticipate being bailed out. In some cases, the decision not to bail in investors provokes a bank run, creating further distortions and leading to even larger bailouts. We ask what macroprudential policies are useful when bailouts crowd out bail-ins.

* We thank seminar and conference participants at the Federal Reserve Bank of Cleveland, the Federal Reserve Board, de Nederlandsche Bank, Rutgers University, the University of Exeter, the 2016 Oxford Financial Intermediation Theory (OxFIT) Conference, the 2nd Chicago Financial Institutions Conference, the Spring 2017 Midwest Macroeconomics Meetings, and especially Fabio Castiglionesi, Huberto Ennis, Charles Kahn and Joel Shapiro for helpful comments. The author thanks the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for research support through grant CRC TR 224.

1 Introduction

In the years since the financial crisis of 2008 and the associated bailouts of banks and other 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 to both protect taxpayers and change the incentives of banks and investors in a way that makes 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 bank itself. This fact raises the question of what incentives banks will face when deciding whether and when to take actions that bail in their investors. In this paper, we ask how the prospect of being bailed out by the government influences banks' 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 bank run, investors fear that a combination of real losses and/or heavy withdrawals will leave their bank 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 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

[1] This 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.

approach.[2] 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 that study more flexible banking arrangements 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.[3]

If the fundamental problem underlying the fragility of banking arrangements is incompleteness of contracts, then an important goal of financial stability policy should be to remove this incompleteness. In other words, a key conclusion of the literature to date is that policy makers 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, certain types of 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.[4]

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

[2] An important exception is Calomiris and Kahn (1991), in which the ex post misallocation of resources associated with a run is part of a desirable ex ante incentive arrangement to disciple bankers' behavior. See also Diamond and Rajan (2001).

[3] This 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 (2010), 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.

[4] See Ennis (2012) for a discussion of the issues involved in reforming money market mutual funds. There is a growing theoretical literature on bail-ins that we do not survey here; see, for example, Walther and White (2017).

contractual environment have been completely successful. We ask whether and under what conditions bank runs can occur in this idealized environment.

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 aim to equalize the marginal utility of consumption across agents to the extent possible.

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. The problems of bank runs and delayed bail-ins can only arise in this model if the underlying losses are sufficiently widespread.

We then analyze possible policy responses to the inefficiencies that arise in the competitive equilibrium. Eliminating bailouts – if possible – would lead banks to immediately bail in their investors when facing losses and would prevent bank runs from occurring in equilibrium. However, it would also eliminate a valuable source of risk sharing and will often lower welfare. We study two policies that can always be used to increase welfare: placing a binding cap on the early payments made by banks and raising additional tax revenue in period 0. We show that the optimal policy combines both of these tools.

The remainder of the paper is organized as follows. The next section describes the economic environment and the actions available to banks, investors, and the government. In Section 3, we derive the constrained efficient allocation of resources in this environment, which is a useful benchmark for what follows. We provide the analysis of equilibrium, including delayed bail-ins and the potential for bank runs, in Section 4. We then discuss possible policy responses in Section 5 before concluding in Section 6.

2 The model

We base our analysis on a version of the Diamond and Dybvig (1983) model with flexible banking contracts and fiscal policy conducted by a government with limited commitment. We introduce idiosyncratic risk to banks' asset holdings and highlight how banks' incentives to react to a loss are influenced by their anticipation of government intervention. In this section, we introduce the agents, preferences, and technologies that characterize the economic environment.

2.1 The environment

Time. There are three time periods, labeled $t = 0, 1, 2$.

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

$$U\left(c_1^{i,k}, c_2^{i,k}, g; \omega^{i,k}\right) \equiv u(c_1^{i,k} + \omega^{i,k} c_2^{i,k}) + v(g), \tag{1}$$

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 π. 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 everywhere greater than one. Each investor is endowed with one unit of of an all-purpose good at the beginning of period 0 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*.[5] Each bank accepts deposits in period 0 from investors in its location and invests these funds 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, $\sigma_k \in \Sigma \equiv \{0, \bar{\sigma}\}$ 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 σ_k as the *fundamental state* of bank k. A bank with $\sigma_k = 0$ is said to have *sound* fundamentals, whereas a bank with $\sigma_k = \bar{\sigma}$ is said to have *weak* fundamentals. The realization of σ_k is observed at the beginning of $t = 1$ by the bank's investors, but is not observed by anyone outside of location k.

After investors' preference types and banks' fundamental states are realized, each investor informs her bank whether she wants to withdraw in period 1 or in period 2. The bank observes *all* reports from its investors *before* making any payments to withdrawing investors. Those investors who chose to withdraw in period 1 then begin arriving sequentially at the bank in a randomly-determined order. Investors are isolated from either other during this process and no trade can occur among them; each investor simply consumes the payment she receives from her bank and returns to isolation. As in Wallace (1988) and others, this

[5] While we use the term "bank" for simplicity, our model should be interpreted as applying to the broad range of financial institutions that engage in maturity transformation.

assumption prevents re-trading opportunities from undermining banks' ability to provide liquidity insurance.

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, all banks have sound fundamentals. In the bad state, in contrast, a fraction $n \in [0, 1]$ of banks have weak fundamentals and, hence, total losses in the financial system are $n\bar{\sigma}$. 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 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 τ. 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, as a best response to the situation at hand.

The government observes the aggregate state of the economy at the beginning of period but, when the aggregate state is bad, is initially unable to determine *which* banks have weak fundamentals. After a measure $\theta \geq 0$ of investors have withdrawn from each bank, the government observes the idiosyncratic state σ_k of all banks and decides how to allocate its tax revenue between bailout payments to banks and the public good. The parameter θ thus measures how quickly the government can collect bank-specific information during a crisis and respond to this information. 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. Once the public good has been provided, the government no longer has access to any resources and there will be no further bailouts.

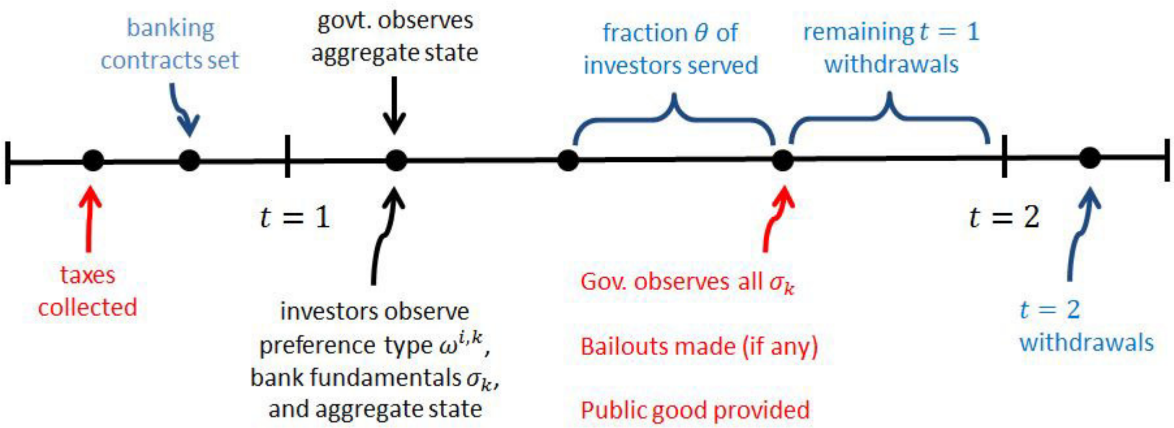


Figure 1: Timeline

2.2 Timeline

The sequence of events is depicted in Figure 1. In period 0, the government chooses the tax rate τ on endowments and investors deposit their after-tax endowment with the bank in their location. At the beginning of period 1, each investor observes her own preference type and the fundamental state of her bank; she then decides whether to withdraw in period 1 or in period 2. Banks observe the choices of their investors and begin making payments to withdrawing investors as they arrive. Once the measure of withdrawals reaches θ, the government observes all banks' fundamental states. At this point, the government may choose to bail out banks with weak fundamentals and places any banks that were bailed out into resolution. After bailout payments are made, all remaining tax revenue is used to provide the public good. Banks that were not bailed out out continue to make payments to investors according to their contract, while the remaining payments made by banks in resolution are dictated by the government.

2.3 Discussion

Sequential service. While our model contains many elements that are familiar from the literature on bank runs, there are some key differences. Perhaps most importantly, banks in our model are able to condition payments to *all* investors on the *total* demand for early withdrawal. Green and Lin (2003) refer to this assumption as "the case without sequential service." This language is potentially confusing when applied to our model: banks still serve withdrawing investors sequentially here. The key point, however, is that a bank is able to observe early withdrawal demand *before* deciding how to allocate resources across agents. By allowing all payments made by the bank to depend on this information, our contract space is larger than that in most of the bank runs literature. In taking this approach, we aim to capture a contractual environment that is sufficiently rich to eliminate the underlying sources of bank runs that appear in the existing literature.

Costly public insurance. The role of aggregate uncertainty in our model is to force the government to fix a tax plan before knowing the aggregate losses of the banking system. If the government knew in advance how many banks would experience loses, it would collect additional taxes at $t = 0$ for the purpose of providing insurance against this location-specific shock. In fact, given that we assume the government can costlessly raise revenue through lump-sum taxes, it would collect enough revenue to provide complete insurance. Our timing assumption makes providing this insurance costly. If, for example, the probability q of the bad state is close to zero, the government will collect tax revenue equal to the desired level of the public good in the good aggregate state. If the realized state turns out to be bad, the marginal value of public resources will increase, but the government will be unable to raise additional revenue.

Delayed intervention. The assumption that the government observes bank-specific information with a delay is important for our analysis because it implies that some investors can withdraw before the government acts. One can narrowly interpret the parameter θ as measuring the time required to both carry out detailed examinations of banks and implement the legal procedures associated with resolving an insolvent bank. More broadly, however, θ can be thought of as also including a variety of other forces that lead governments to act slowly in the early stages of a crisis. For example, investors who are well-connected politically may use their influence to delay any government intervention until after they have had an opportunity to withdraw. The timing of the intervention might also reflect opaque incentives faced by regulators.[6] In addition, Brown and Dinc (2005) provide evidence that the timing of a government's intervention in resolving a failed financial institution depends on the electoral cycle. Looking at episodes from 21 major emerging market economies in the 1990s, they find that interventions that would impose large costs on taxpayers and/or would more fully reveal the extent of the crisis were significantly less likely to occur before elections. (See also Rogoff and Sibert, 1988.) The effect of such political factors that delay the policy response

[6] Kroszner and Strahan (1996) argue that throughout the 1980s the Federal Savings and Loan Insurance Corporation (FSLIC) faced 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.

to a crisis would be captured in our model by an increase in the parameter θ.

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, as well as investors' withdrawal decisions. This planner observes all of the information available to banks and investors, including each investor's preference type. It faces the same restrictions on fiscal policy as the government; in particular, all tax revenue must be raised at $t = 0$, before the aggregate state is realized. The planner allocates resources to maximize the sum of all investors' utilities.

It is fairly easy to see that the planner will direct all impatient investors to withdraw at $t = 1$, since they do not value later consumption, and will direct all patient investors to withdraw at $t = 2$, since it is less expensive to provide consumption to them after investment has matured. In addition, because investors are risk averse, the planner will choose to treat investors and banks symmetrically. In the good aggregate state, the planner will give a common level of consumption c_{10} in period 1 to all impatient investors and a common level c_{20} in period 2 to all patient investors. (The second subscript indicates that these consumption levels pertain to the good aggregate state, where zero banks have weak fundamentals.) In the bad aggregate state, the planner will give a common consumption profile (c_{1S}, c_{2S}) to investors in all banks with strong fundamentals and a common profile $(c_{1W}, _{2W})$ to investors in all banks with weak fundamentals. These consumption levels will be chosen to maximize

$$(1-q)\left\{\pi u\left(c_{10}\right)+(1-\pi)\,u\left(c_{20}\right)+v\left(\tau\right)\right\}$$
$$+q\left\{\begin{array}{c}(1-n)\left(\pi u\left(c_{1S}\right)+(1-\pi)\,u\left(c_{2S}\right)\right)+n\left(\pi u\left(c_{1W}\right)+(1-\pi)\,u\left(c_{2W}\right)\right)\\+v\left(\tau-(1-n)b_S-nb_W\right)\end{array}\right\}.$$

subject to feasibility constraints

$$\pi c_{10}+(1-\pi)\frac{c_{20}}{R}\leq 1-\tau \tag{2}$$
$$\pi c_{1S}+(1-\pi)\frac{c_{2S}}{R}\leq 1-\tau+b_S \tag{3}$$
$$\pi c_{1W}+(1-\pi)\frac{c_{2W}}{R}\leq 1-\tau-\bar{\sigma}+b_W, \tag{4}$$

where b_z denotes the per-investor transfer (or "bailout") given to each bank of type z in the bad aggregate state. These constraints each state that the present value of the consumption given to depositors in a bank must come from the initial deposit $1-\tau$, minus the loss $\bar{\sigma}$ for banks with weak fundamentals, plus any bailout received.[7] The restriction that the planner cannot raise additional tax revenue in period 1 is equivalent to saying that the bailout payments must be non-negative,

$$b_S \geq 0 \quad \text{and} \quad b_W \geq 0. \tag{5}$$

The first-order conditions for the optimal consumption levels can be written as

$$u'(c_{1z}) = Ru'(c_{2z}) = \mu_z \qquad \text{for } z = 0, S, W, \tag{6}$$

where μ_z is the Lagrange multiplier on the resource constraint associated with state z normalized by the probability of a bank ending up in that state. The first-order condition for the choice of tax rate τ can be written as

$$(1-q)v'(\tau) + qv'(\tau - (1-n)b_S - nb_W) = (1-q)\mu_0 + q(1-n)\mu_S + qn\mu_W, \tag{7}$$

which states that the expected marginal value of a unit of public consumption equals the expected marginal value of a unit of private consumption at $t = 0$. The first-order conditions for the bailout payments are

$$v'(\tau - (1-n)b_S - nb_W) \geq \mu_z, \qquad \text{with equality if } b_z > 0, \qquad \text{for } z = S, W. \tag{8}$$

If the marginal value of private consumption in some banks were higher than the marginal value of public consumption in the bad aggregate state, the planner would transfer resources to (or "bail out") these banks until these marginal are equalized. If instead the marginal value of private consumption in a bank is lower than the marginal value of public consumption, the bank will not be bailed out and the constraint in (5) will bind.

The following two propositions characterize the key features of the constrained efficient allocation of resources in our environment. First, the consumption of investors in banks with sound fundamentals is independent of the aggregate state and these banks do not receive bailouts.[8]

[7] Note that our notation does not allow the planner to make bailout payments in the good aggregate state. This assumption prevents the planner from being able to make tax revenue fully state-contingent by, for example, setting $\tau = 1$ and holding all resources outside of the banking system until the aggregate state is revealed.

[8] The first part of this result depends on our simplifying assumption that sound banks are completely unaffected by the bad aggregate state, but the second part of the result does not. Even if sound banks were to experience some losses during an economic downturn, the planner would not choose to bail out these banks as long as the losses are small relative to those at weak banks.

Proposition 1. *The constrained efficient allocation satisfies*

$$(c_{10}^*, c_{20}^*) = (c_{1S}^*, c_{2S}^*) \quad \textit{and} \quad b_S^* = 0.$$

Given this result, we will drop the (c_{10}, c_{20}) notation in what follows and use (c_{1S}, c_{2S}) to refer to the consumption profile for investors in a bank with sound fundamentals regardless of the aggregate state. Our second result shows that this profile is different from the one assigned by the planner to investors in banks with weak fundamentals.

Proposition 2. *The constrained efficient allocation satisfies*

$$(c_{1S}^*, c_{2S}^*) \gg (c_{1W}^*, c_{2W}^*) \quad \textit{and} \quad b_W^* > 0.$$

This result shows that the constrained efficient involves a combination of *bailouts* and *bail-ins* for investors in banks with weak fundamentals. The optimal bailout b_W^* gives investors partial insurance against the risk associated with their bank's losses. However, 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.

It is worth pointing out that the constrained-efficient bail-in applies equally to *all* investors in a weak bank, regardless of when they arrive to withdraw. While the desirability of this feature follows immediately from risk aversion, we will see below that it often fails to hold in a decentralized equilibrium. It is also worth noting that the constrained efficient allocation is incentive compatible. The first-order conditions (6) and $R > 1$ imply that $c_{1z}^* < c_{2z}^*$ holds for every state z and, hence, a patient investor always prefers her allocation to that given to an impatient investor (and vice versa).

4 Equilibrium

In this section we begin our investigation of the decentralized economy. Compared to the planner's economy discussed in the previous section, the decentralized economy differs in the following important ways. First, investors' preference types are private information and the banking contract therefore allows investors to choose the period in which they withdraw. Second, each bank is concerned solely with its own investors and takes economy-wide variables, including the level of the public good, as given. Third, there is asymmetric information between the banks and the government; while the government immediately observes the aggregate state at the beginning of $t = 1$, it must wait for θ withdrawals to take place before observing bank-specific states. The government then makes bailout payments to banks with weak fundamentals and places these bank into resolution. Importantly, the bailout and resolution policies cannot be set ex-ante, but instead are chosen as a best response to the situation at hand.

In this section, we study equilibrium in the withdrawal game played by an individual bank's investors, taking the actions of investors at other banks (and the government) as given. In section 5, we study the joint determination of equilibrium actions across all banks.

4.1 Preliminaries

We begin by reviewing the timeline of events in Figure 1 for the decentralized economy and then provide a general definition of equilibrium.

The tax rate. To simplify the analysis in this section, we assume that the tax rate τ levied by the government in period 0 is set to the value from the constrained efficient allocation, τ^*. We derive equilibrium withdrawal behavior and the equilibrium allocation of resources for this given tax rate. In Section 5, we examine the government's optimal choice of tax rate given the equilibrium outcomes identified in this section.

Banking contracts. In period 0, each bank establishes a contract that specifies how much it will pay to each withdrawing investor as a function of both the bank's fundamental state $\sigma_k \in \{0, \bar{\sigma}\}$ and the fraction $\rho^k \in [\pi, 1]$ of its investors who choose to withdraw early. We allow the government to set an upper bound $\bar{c}$ on the payments made to any investor withdrawing in period 1. One way to justify this upper bound is to assume that while the government cannot dictate the exact terms of the contractual arrangement between a bank and its investors, it is able to impose broad guidelines on the types of contract banks are allowed to offer.

Because investors are risk averse, it will be optimal for a bank to give the same level of consumption to all investors who withdraw in the same period.[9] Let c_1^k denote the payment made by the bank to each investor who withdraws in period 1. In period 2, the bank

[9] Keep in mind that our environment is different from that studied by Wallace (1990), Green and Lin (2003), Peck and Shell (2003) and others where the bank gradually learns about the demand for early withdrawal by observing investors' actions as they take place. Here, a bank directly observes total early withdrawal demand before making any payments to investors. It learns no new information as investors sequentially withdraw at $t = 1$ and, therefore, an optimal arrangement will always give the same level of consumption to each of these investors.

will divide its matured investment, plus any bailout payment received, evenly among its remaining depositors. The operation of the bank is, therefore, completely described by the function

$$c_1^k : \{0, \bar{\sigma}\} \times [\pi, 1] \to [0, \bar{c}]. \tag{9}$$

We refer to the function in (9) as the *banking contract*. There is full commitment to the banking contract in the sense that the plan in (9) will be followed unless the bank is placed into resolution by the government. Each bank's contract is chosen to maximize the expected utility from private consumption of the bank's investors.[10]

Bailouts and resolution. After a fraction θ of investors have withdrawn at $t = 1$, the government observes the fundamental state σ_k of each bank and chooses a bailout payment b^k for each bank with weak fundamentals. It then dictates the payments made by these banks to their remaining investors as part of the resolution process. We characterize the government's bailout/resolution policy below.

Withdrawal strategies. An investor's withdrawal decision can depend on both her preference type ω_i^k and the fundamental state of her bank σ_k. (See Figure 1.) A withdrawal strategy for investor i in bank k is, therefore, a mapping:

$$y_k^i : \Omega \times \Sigma \to \{0, 1\}$$

where $y_k^i = 0$ corresponds to withdrawing in period 1 and $y_k^i = 1$ corresponds to withdrawing in period 2. An investor will always choose to withdraw in period 1 if she is impatient. We introduce the following labels to describe the actions an investor takes in the event she is patient.

Definition 1. For given σ_k, we say investor i in bank k follows:
(i) the *no-run strategy* if $y_k^i(\omega_k^i, \sigma_k) = \omega_k^i$ for $\omega_k^i \in \{0, 1\}$, and
(ii) the *run strategy* if $y_k^i(\omega_k^i, \sigma_k) = 0$ for $\omega_k^i \in \{0, 1\}$.

We use y_k to denote the profile of withdrawal strategies for all investors in bank k and y to denote the withdrawal strategies of all investors in the economy. It will often be useful to summarize a profile of withdrawal strategies by the fraction of investors who follow the run

[10] This outcome would obtain, for example, if multiple banks competed for deposits in each location. We use a representative bank in each location only to simplify the presentation.

strategy in that profile, which we denote

$$x_{\sigma_k} \equiv \int_{i\in[0,1]} (1 - y_k^i(\omega_k^i = 1, \sigma_k))dk.$$

Similarly, we use ρ_k to denote the total demand for early withdrawal from bank k in a given profile, which equals

$$\rho_k = \pi + (1-\pi)x_k.$$

Allocations. The allocation of private consumption in bank k in a particular state is a specification of how many investors withdraw at $t = 1$ in that state, how much consumption each of these investors receives, and how much consumption each remaining investor receives at $t = 2$. This allocation depends on the banking contract for bank k, the withdrawal strategies of investors in bank k, and the government intervention in bank k (if any). The details of the government intervention, in turn, may depend on the contracts of other banks and the withdrawal strategies of investors in those banks. In general, therefore, the optimal withdrawal behavior for each investor in bank k may depend on the contracts offered by *other* banks and on the withdrawal strategies of investors in *other* banks.

Equilibrium. To study equilibrium withdrawal behavior within a single bank, we fix all banking contracts, the government's intervention policy, and the withdrawal strategies of investors in all other banks, y_{-k}. Together, these items determine the payoffs of what we call the *withdrawal game in bank k*. That is, holding these other items fixed, we can calculate the allocation of private consumption in bank k as a function of the strategies y_k played by that bank's investors. An equilibrium of this game is a profile of strategies for the bank's investors, y_k, such that for every investor i in the bank, y_k^i is a best response to the strategies of the other investors, y_k^{-i}.

An equilibrium of the overall economy is a profile of withdrawal strategies for all investors y^* such that (i) y_k^* is an equilibrium of the withdrawal game in bank k generated by the strategies y_{-k}^* of investors in all other banks, for all k, (ii) the contract in bank k maximizes the expected utility of its investors taking as given the contracts and withdrawal strategies y_{-k}^* of investors in all other banks, for all k, and (iii) the government's bailout and resolution policy maximizes total welfare taking as given all banking contracts and withdrawal strategies y^*. Notice how this definition reflects the timing assumptions depicted in Figure 1. Investors in bank k recognize that their choice of contract will influence equilibrium withdrawal behavior within their own bank but will not affect outcomes at other banks.[11] The

[11] This result follows, in part, from the assumption that there are a continuum of locations and, hence, the

government's bailout and resolution policies, in contrast, are set after all banking contracts and withdrawal decisions have been made. Because the government cannot commit to the details of these policies ex ante, it acts to maximize welfare taking all bank contracts and withdrawal decisions as given.

In the subsections that follow, we derive the properties of the contracts that a bank will use in equilibrium, focusing first on the case where its fundamental state is sound. We then turn to the case where the bank's fundamental state is weak, which requires characterizing the optimal bailout and resolution policies as well. Finally, we then characterize the equilibrium of the entire economy, in which each bank's contract is a best response to all other contracts.

4.2 Banks with sound fundamentals

We assume the government does not give bailouts to banks with sound fundamentals, nor does it place them in resolution.[12] As a result, all investors who chose to withdraw at $t = 1$ receive the contractual amount $c_1^k(0,\, \rho^k)$, and all investors who chose with withdraw at $t = 2$ receive an even share of the bank's assets, $c_2^k(0,\, \rho^k)$, which is implicitly defined by

$$p^k c_1(0,\, \rho^k) + (1 - \rho^k)\frac{c_2(0,\, \rho^k)}{R} = 1 - \tau. \tag{10}$$

The bank and its investors recognize that ρ^k will result from the equilibrium withdrawal behavior of investors. In particular, if the bank offers a higher payment in period 1 than in period 2, all investors will choose to withdraw early. In other words, equilibrium requires

$$\rho^k = \left\{ \begin{array}{c} \pi \\ \in [\pi, 1] \\ 1 \end{array} \right\} \quad \text{as} \quad c_1(0,\, \rho^k) \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} c_2(0,\, \rho^k). \tag{11}$$

We refer to (11) as the *implementability constraint*. If a triple (ρ_S, c_{1S}, c_{2S}) satisfy both (10) and (11), then any banking contract with $c_1^k(\sigma_k, \rho_S) = c_{1S}$ for $\sigma_k = 0$ can implement this allocation as an equilibrium of the withdrawal game in bank k, regardless of how the payments $c_1^k(\sigma_k, \rho^k)$ are set for other values of ρ^k. The following result shows that something stronger is true: by choosing these other payments appropriately, the banking contract can be set so that the withdrawal game in bank k has a *unique* equilibrium.

---

actions taken at one bank have no effect on aggregate variables or on the behavior of the government toward other banks.

[12] Recall from Section 3 that the constrained efficient allocation involves zero bailouts for sound banks. Our assumption here is that the government is able to commit to follow this policy.

Proposition 3. *If the allocation* (ρ_S, c_{1S}, c_{2S}) *satisfies both* (10) *and* (11)*, there exists a contract that implements this allocation as the unique equilibrium of the withdrawal game played by a sound bank's investors.*

In light of the above proposition, we can recast the problem of choosing the optimal banking contract as one of directly choosing the allocation (ρ_S, c_{1S}, c_{2S}) to maximize expected utility

$$V_S(\rho_S,\, c_{1S}) \equiv \rho_S u(c_{1S}) + (1-\rho_S)u(c_{2S}) \tag{12}$$

subject to the feasibility constraint (10) and the implementability constraint (11) and the restriction $c_1^k \in [0, \bar{c}]$ for all banks. The next result characterizes the solution to this problem.

Proposition 4. *When bank k has sound fundamentals, there is a unique equilibrium of the withdrawal game in the bank associated with the optimal banking contract. The equilibrium allocation* (ρ_S, c_{1S}, c_{2S}) *satisfies* $\rho_S = \pi$ *and* $c_{1S} = min\{c_{1S}^*, \bar{c}\}$.

This result shows that as long as the upper bound $\bar{c}$ is set high enough to allow it, the equilibrium allocation within a sound bank is the same as in the constrained efficient allocation. In other words, resources are always allocated efficiently within a sound bank and investors never run on these banks.

There are many banking contracts that implement the desired allocation, one of which is

$$c_1^k\left(0,\, \rho^k\right) = \left\{ \begin{array}{c} min\{c_{1S}^*, \bar{c}\} \\ 0 \end{array} \right\} \quad \text{if} \quad \left\{ \begin{array}{c} \rho^k = \pi \\ \rho^k > \pi \end{array} \right\}. \tag{13}$$

Under this contract, the bank would immediately suspend withdrawals if more than a fraction π of its investors request to withdraw early, saving all of its resources until period 2.[13] It is easy to verify that waiting to withdraw in period 2 is then the best response for a patient investor to any profile of withdrawal strategies for the other investors. As a result, this contract uniquely implements the desired allocation in the withdrawal game in bank k and a bank run will never occur.

4.3 Banks with weak fundamentals

Characterizing the outcome of the withdrawal game in a weak bank is more complicated because it depends on the government's bailout and resolution policies. Let W denote the set of weak banks,

[13] The reaction of setting early payments to zero when $\rho_k > \pi$ holds is stronger than needed to eliminate the run equilibrium. It would suffice to set these payments low enough that a patient investor would receive more consumption by waiting to withdraw in period 2. The key point is that the bank can easily structure the contract to prevent a run; the form in equation (10) makes this point in a particularly clean way.

$$W \equiv \{k \in [0,1] \text{ s.t } \sigma_k = \bar{\sigma}\}\,. \tag{14}$$

After the first θ withdrawals have taken place in all banks, the government observes the fundamental state σ_k of each bank. For $k \in W$, the government also observes the bank's current condition: the amount of resources remaining in the bank and the fraction of the bank's remaining investors who are impatient. The government then decides on a bailout payment b^k for each $k \in W$ and places these banks into resolution. We derive the government's best responses by working backward, beginning with the resolution stage.

Resolution. To simplify the presentation, we assume that when a bank is placed into resolution, the government directly observes the preference types of the bank's remaining investors and allocates the bank's resources (including the bailout payment) conditional on these types. One could imagine, for example, the using the court system to evaluate individual's true liquidity needs, as discussed in Ennis and Keister (2009). This assumption is **not** important for our results, however. If instead the government were to offer a new banking contract and have the remaining investors play a withdrawal game based on this new contract, it could choose a contract that yields the outcome we study here as the unique equilibrium of that game.

Let $\hat{\psi}_k$ denote the per-capita level of resources in bank k, including any bailout payment received, after the first θ withdrawals have taken place. Then we have

$$\hat{\psi}_k = \frac{(1-\tau)(1-\bar{\sigma}) - \theta c_1^k(\bar{\sigma},\, \rho_k) + b_k}{1-\theta}, \tag{15}$$

where b_k is the per-investor bailout given to bank k. Let $\hat{\rho}^k$ denote the fraction of the bank's remaining investors who are impatient. The allocation of resources for a bank in resolution is chosen to maximize the sum of the utilities for the remaining investors in the bank:

$$\hat{V}\left(\hat{\psi}_k;\, \hat{\rho}_k\right) \equiv \max_{\hat{c}_1^k,\, \hat{c}_2^k}\ (1-\theta)\left(\hat{\rho}_k u\left(\hat{c}_1^k\right) + (1-\hat{\rho}_k) u\left(\hat{c}_2^k\right)\right) \tag{16}$$

subject to the feasibility constraint

$$\hat{\rho}_k \hat{c}_1^k + (1-\hat{\rho}_k)\frac{\hat{c}_2^k}{R} \le \hat{\psi}_k \tag{17}$$

The optimal choice of post-bailout payments is determined by the first order condition

$$u'\left(\hat{c}_1^k\right) = Ru'\left(\hat{c}_2^k\right) = \hat{\mu}(\hat{\psi}_k;\, \hat{\rho}_k), \tag{18}$$

where $\hat{\mu}$ is the Lagrange multiplier on the resource constraint. Since $R > 1$, this condition implies that a bank in resolution provides more consumption to patient investors withdrawing in period 2 than to the remaining impatient investors who withdraw in period 1.

Bailouts. In choosing the bailout payments $\left\{b^k\right\}$, the government's objective is to maximize the sum of the utilities of all investors in the economy. While bailouts raise the private consumption of investors in weak banks, they lower the provision of the public good, which affects all investors. The government' objective in choosing these payments can be written as

$$\max_{\left\{b^k\right\}_{k\in W}} \int_W \hat{V}\left(\hat{\psi}_k;\, \hat{\rho}_k\right) dk + v\left(\tau - \int_W \hat{b}_k dk\right) \tag{19}$$

The first-order condition for this problem is

$$\hat{\mu}(\hat{\psi}_k;\, \hat{\rho}_k) = v'\left(\tau - \int_W \hat{b}_k dk\right) \text{ for all } k. \tag{20}$$

Notice that the right-hand side of this equation – the marginal utility of public consumption – is independent of k. The optimal bailout policy thus has the feature that the marginal value of resources will be equalized across all weak banks, regardless of their chosen banking contract or the withdrawal behavior of their investors. As a result, all banks in resolution will give a common consumption allocation $(\hat{c}_1, \hat{c}_2)$ to their impatient and patient investors, respectively. These consumption values and the bailout payments b^k will satisfy the resource constraint

$$\hat{\rho}_k \hat{c}_1 + (1 - \hat{\rho}_k)\frac{\hat{c}_2}{R} = \frac{(1-\tau)(1-\bar{\sigma}) - \theta c_1^k\left(\bar{\sigma},\, \rho_k\right) + b_k}{1-\theta}. \tag{21}$$

Using the fact that $(\hat{c}_1, \hat{c}_2)$ is the same in all weak banks, this constraint shows that the bailout payment made to bank k is increasing in the amount paid out by the bank before being bailed out, $c_1^k\left(\bar{\sigma},\, \rho_k\right)$. Together with the first-order condition (18), this constraint implies that b^k is increasing in the fraction of bank k's remaining investors who are impatient, $\hat{\rho}_k$.

Withdrawal behavior. A fraction θ of a weak bank's investors will receive the amount specified by the contract, $c_1^k\left(\bar{\sigma},\, \rho_k\right)$, before the government intervenes. The bank will then be bailed out and placed into resolution. Its remaining impatient investors will receive $\hat{c}_1$ and its remaining patient investors will receive $\hat{c}_2$, as derived above. A patient investor will choose to withdraw early if the contract sets $c_1^k > \hat{c}_2$ and will choose to wait if $c_1^k < \hat{c}_2$. In other words, the fraction ρ^k of investors who attempt to withdraw from a weak bank at $t = 1$ will satisfy

$$\rho^k = \begin{Bmatrix} \pi \\ \in [\pi, 1] \\ 1 \end{Bmatrix} \quad \text{if} \quad c_1^k(\bar{\sigma}, \rho_k) \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \hat{c}_2. \tag{22}$$

In choosing a contract, the bank recognizes that its investors will behave in accordance with (22), which we refer to as the implementability constraint for weak banks. The next result is the analog of Proposition 3 from the previous section: it shows that any allocation satisfying the implementability constraint can be implemented as the unique equilibrium of the withdrawal game in bank k.

Proposition 5. *If* (ρ_W, c_{1W}) *satisfy* (22)*, then there exists a banking contract* c_1^k *that implements this allocation as the unique equilibrium of the withdrawal game played by a weak bank's investors.*

This results allows us to formulate the bank's optimal contract problem as one of directly choosing the allocation (ρ_W, c_{1W}) to maximize

$$V_W(\rho_W, c_{1W}) \equiv \theta u(c_{1W}) + (1-\theta)\left[\hat{\rho}_{\bar{\sigma}} u(\hat{c}_1) + (1-\hat{\rho}_{\bar{\sigma}}) u(\hat{c}_2)\right] \tag{23}$$

subject to the implementability constraint for weak banks (22) and the relationship

$$\hat{\rho}_{\bar{\sigma}} \equiv \frac{\pi}{1-\theta}\left(\frac{\rho_W - \theta}{\rho_W}\right). \tag{24}$$

This last expression shows how the fraction of the bank's remaining investors after θ withdrawals are impatient depends on the fraction that initially attempt to withdraw early.

The first term in the objective function in (23) is clearly increasing in the choice of c_{1W}, reflecting the bank's desire to give as much consumption as possible to the investors who withdraw before the bank is placed into resolution. However, the implementability constraint (22) shows that if c_{1W} is set greater than $\hat{c}_2$, the bank's investors will run, in which case ρ^k will equal 1. A run on the bank is costly because some early consumption is inefficiently given to patient investors; the fact can be seen by noting that $\hat{\rho}_{\bar{\sigma}}$ is an increasing function of ρ^k and the second term in the objective function (23) is strictly decreasing in $\hat{\rho}_{\bar{\sigma}}$. The next result shows that the solution to the bank's problem takes one of two forms: the early payments will either be set as high as possible, or will be set to the largest value that prevents a bank run.

Proposition 6. *The solution to the program of maximizing* (23) *subject to* (22) *and* (24) *will either set* $c_{1W} = \bar{c}$ *or* $c_{1W} = \hat{c}_2$*.*

Which of these two options will be optimal for the bank depends on how low the early payment would need to be set in order to prevent a run. Setting c_{1W} equal to the upper bound $\bar{c}$ will be optimal whenever

$$V_W(\bar{c},\, 1) > V_W(\hat{c}_2,\, \pi)\,. \tag{25}$$

The above inequality implies that the loss to the remaining $1-\theta$ investors in the bank resulting from keeping payments as high as possible and allowing a run is more than offset by the gain to the first fraction θ to withdraw. By observing that the inequality in (25) is equivalent to

$$u(\bar{c}) - u(\hat{c}_2) > (1-\pi)\left(u(\hat{c}_2) - u(\hat{c}_1)\right)$$

the equilibrium of the withdrawal game in a weak bank can be characterized as follows.

Proposition 7. *If bank k has weak fundamentals then:*

(i) If $u(\bar{c})-u(\hat{c}_2) < (1-\pi)\left(u(\hat{c}_2) - u(\hat{c}_1)\right)$, there is a unique equilibrium of the withdrawal game in bank k associated with the optimal banking contract. The equilibrium allocation has $\rho_W = \pi$ and $c_{1W} = \min\{\bar{c},\, \hat{c}_2\}$.

(ii) If $u(\bar{c}) - u(\hat{c}_2) > (1-\pi)\left(u(\hat{c}_2) - u(\hat{c}_1)\right)$, there is again a unique equilibrium of the withdrawal game in bank k associated with the optimal banking contract. The equilibrium allocation in this case has $\rho_W = 1$ and $c_{1W} = \bar{c}$.

(iii) If $u(\bar{c}) - u(\hat{c}_2) = (1-\pi)\left(u(\hat{c}_2) - u(\hat{c}_1)\right)$, the withdrawal game in bank k has multiple equilibria, one with $\rho_W = \pi$ and $c_{1W} = \hat{c}_2$ and another with $\rho_W = 1$ and $c_{1W} = \bar{c}$.

This result shows that, outside of a knife-edge case, the withdrawal game in a weak bank will have a unique equilibrium under the optimal banking contract. In this sense, a bank run in our model is fundamentally different from the type of self-fulfilling run normally studied in the literature based on Diamond and Dybvig (1983). When a bank is in case *(ii)* of Proposition 7, withdrawing early is a dominant strategy for the bank's investors. In this sense, a bank run in our model does not rely on investors' self-fulfilling beliefs about the actions of other investors in their bank. We show below, however, that the model may still have multiple equilibria because an investor's best response may depend critically on the withdrawal decisions of investors in *other* banks.

4.4 Equilibrium across banks

The preceding sections have investigated the equilibrium outcomes within a given bank, taking the actions of the government and the remaining banks as fixed. We now investigate the properties of the overall equilibrium across all banks, in which both the banking contract and the withdrawal strategies in each bank are best responses to the actions taking place at other banks.

Constrained inefficiency. We begin by asking whether the equilibrium allocation is constrained efficient. Note that, in order for this allocation to be feasible in the decentralized economy, the upper bound $\bar{c}$ on early payments must be set sufficiently high that sound banks are able to choose c_{1S}^*. For the analysis in this section, we will set $\bar{c} = c_{1S}^*$. Our next result shows that, even though it is feasible, the constrained efficient allocation is never an equilibrium of the decentralized economy.

Proposition 8. *The equilibrium allocation of resources is never constrained efficient.*

The bailout policy creates an incentive for weak banks to set their early payments as high as possible. The only reason a weak bank would voluntarily impose losses on its investors (by setting a payment below $\bar{c} = c_{1S}^*$) is to prevent a run. Note that preventing a run only requires that the payment in period 1 not exceed $\hat{c}_2$ and, as a result, a weak bank will never set its early payment below this level. In particular, a weak bank will never choose to bail in its investors all the way down to $\hat{c}_1$, as occurs in the constrained efficient allocation.

Equilibrium bank runs. In addition to being constrained inefficient, the equilibrium of the full model will, in some cases, involve a run by investors on weak banks.

Proposition 9. *For some parameter values, there exists an equilibrium in which investors run on weak banks. In some cases this equilibrium is unique, but in others it coexists with another equilibrium in which no run occurs.*

In the run equilibrium, all investors in weak banks attempt to withdraw at $t = 1$, that is, the profile of withdrawal strategies has $x_{\bar{\sigma}} = 1$. A fraction θ of these investors successfully withdraw before the government observes $\sigma_k = \bar{\sigma}$ and places the bank into resolution.

The result in Proposition 9 is established in Figure 2 which depicts the type of equilibria that arise for different combinations of the parameters n, the fraction of weak banks, and $\bar{\sigma}$, the loss in each of them. The figure uses the utility function[14]

$$u(c_1^{i,k} + \omega^{i,k} c_2^{i,k}) = \frac{\left(c_1^{i,k} + \omega^{i,k} c_2^{i,k}\right)^{1-\gamma} - 1}{1-\gamma} \qquad \text{and} \qquad v(g) = \delta \frac{g^{1-\gamma}}{1-\gamma}. \tag{26}$$

For parameter combinations in the dark region in the lower-left part of the graph, there is a unique equilibrium of the model and the allocation in this equilibrium does not involve a bank

[14] The other parameters of the model are set to $R = 1.5$, $\pi = 0.5$, $\gamma = 5$, $\delta = 0.5$, $q = 0.05$ and $\theta = 0.5$. The tax rate τ is set to its constrained efficient value from section 3.

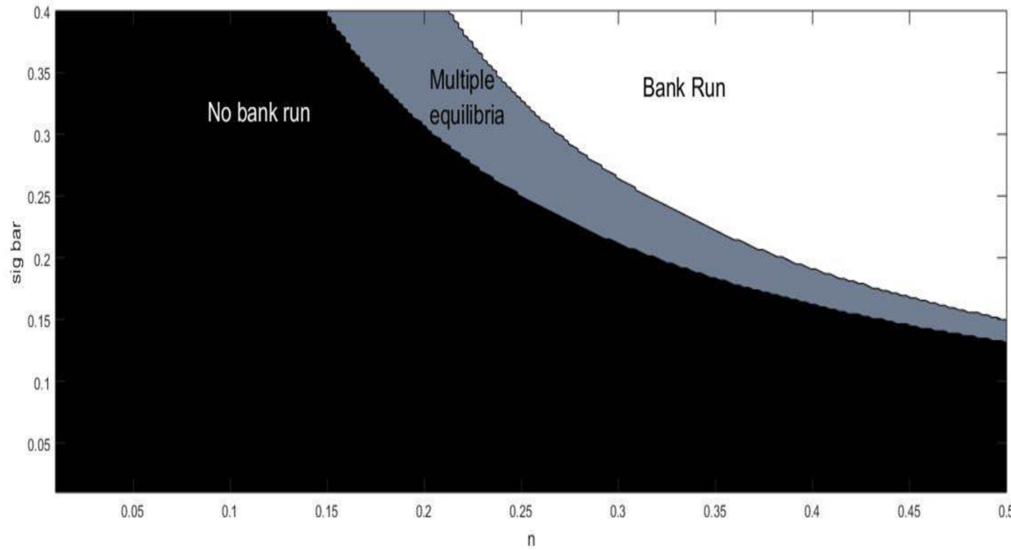


Figure 2: Equilibrium with a bank run

run. When the losses $\bar{\sigma}$ suffered by a weak bank are small and/or few banks experience these losses, the process of resolving these banks has a relatively small cost for the government. When this cost is small, the government remains in good fiscal condition and will choose to make bailout payments that lead to relatively high consumption levels $(\hat{c}_1, \hat{c}_2)$ for the remaining investors in banks placed into resolution. This fact, in turn, makes running in an attempt to withdraw before the government intervenes less attractive for patient investors in a weak bank. As a result, a unique equilibrium exists and all patient investors wait until $t = 2$ to withdraw.

In the unshaded region in the upper-right portion of the figure, in contrast, both the number of banks experiencing a loss and the amount lost by each of these banks are significant. In this case, the government's budget constraint will be substantially impacted by its desire to bail out weak banks in a crisis. As the marginal value of public resources rises, the bailout and resolution process will lead to lower consumption levels $(\hat{c}_1, \hat{c}_2)$ for the remaining investors in these banks. When $\hat{c}_2$ is low enough, the equilibrium within a weak bank k will involve a run by patient investors, as shown in Proposition 7. The overall equilibrium in this region is again unique, but the (larger) losses on weak banks' asset are now compounded by the additional liquidation of assets and misallocation of resources created by the run.

Multiple equilibria. In the grey region in Figure 2, both of the equilibria described above exist. The fact that multiple equilibria exist in this region is particularly interesting in light of Proposition 7, which showed that the equilibrium of the withdrawal game within each bank is unique except for in a knife-edge case. The multiplicity of equilibria illustrated in Figure 2 arises because of an externality in payoffs *across* weak banks. When a run occurs at other weak banks, this event causes more investment to be liquidated and leads to larger bailouts at those banks. The larger bailouts place greater strain on the government's budget constraint and lead – all else being equal – to a smaller bailout at bank k. In the lighter-shaded region in Figure 2, this smaller bailout lowers the consumption levels $(\hat{c}_1, \hat{c}_2)$ enough to make running a best response for the patient investors in bank k. In other words, in our model there is a strategic complementarity in the withdrawal decisions of investors across banks. The usual strategic complementarity that appears in models in the Diamond-Dybvig tradition – which arises between investors within a bank – is eliminated by the more flexible banking contracts. However, the government's bailout and resolution policy introduces this new complementarity in actions across banks, which creates the region of multiple equilibria in Figure 2.

It is worth emphasizing that a run on bank k lowers the welfare of the bank's investors in much the same way as in the existing literature. Holding fixed the bailout payment it receives, a bank's investors would be strictly better off if there were no run on the bank. Moreover, the bank has contractual tools that would allow it to prevent the run. The problem, however, is that preventing the run requires decreasing the payment given to the first θ investors who withdraw, and this action would decrease the bailout payment the bank receives. Instead, in this equilibrium, the bank's investors choose to tolerate the run as a side effect of the plan that maximizes the level of payments the bank is able to make to its investors before the government intervenes.

Runs cannot be based on sunspots alone. The pattern in Figure 2 suggests that the run equilibrium does not exist for $\bar{\sigma}$ equal or sufficiently close to zero. The next proposition shows that this property holds more generally.

Proposition 10. *Given other parameter values, there exists $\bar{\sigma}^* > 0$ such that a bank run does not occur in equilibrium for any $\bar{\sigma} < \bar{\sigma}^*$.*

According to Proposition 10, a run in this environment cannot occur unless the fundamental was large enough. The reason is as follows. In the bad aggregate state, a sound bank would deliver higher utility to its investors compared to a weak bank. At the same time - conditional on weak banks *not* experiencing a run - the utility gain associated with being sound is approaching zero as $\bar{\sigma} \to 0$. If the run equilibrium exists for $\bar{\sigma}$ close to zero, then almost all of the losses in the weak banks will be generated solely by the run from their investors. In this case, an individual weak bank can always do better by deviating from equilibrium play and implementing a contract with strong suspension clause both when the bank is sound and when the bank is weak. Indeed, since $\bar{\sigma} \approx 0$, by preventing a run, bank-k also ensures that its investors receive almost identical welfare as the investors in sound banks which in turn is strictly higher than the welfare in the remaining weak banks (since those weak banks experience a run). This reasoning establishes that the run equilibrium cannot exist for $\bar{\sigma}$ equal or sufficiently close to zero since an individual weak bank strictly gains by implementing a run-proof contract.

Runs must be systemic. Unlike other models of financial fragility, a bank run in our model cannot be an isolated event in the sense of occurring at a single bank in or model or even a small group of banks. The pattern in Figure 2 also suggests that a run can only occur when the number of weak banks is sufficiently large. The following proposition formalizes this result.

Proposition 11. *Given other parameter values, there exists $n^* > 0$ such that a bank run does not occur in equilibrium for any $n < n^*$.*

If the number of affected banks is small, the associated losses will have a small impact on the government's budget constraint. If the government remains in good fiscal condition, the bailout policy it will choose ex post treats weak banks generously, leaving their patient investors with no incentive to run.

Observe how the environment we study is different from most of the literature on banking crisis. Here, runs will not occur in equilibrium unless there were real shocks $\bar{\sigma} > 0$, this shocks were large enough (Proposition 10) and sufficiently wide spread (Proposition 11). An implication of Propositions 10 and 11 is that a run cannot be based on sunspots alone, regardless of how many banks may have experienced the "*bad sunspot*". Finally, it is important to stress that Propositions 10 and 11 are *not* stating that sunspots cannot play a role our framework. Instead, a necessary condition for a sunspot state to affect equilibrium outcomes is that n and $\bar{\sigma}$ must both be positive and to belong to the region of parameters where both the run and the no-run equilibrium co-exist (Proposition 9). In this case, one can introduce a sunspot that that would serve to coordinate investors' withdrawal decisions on one or the other equilibrium. In this sense, a bank run would still be *self-fulfilling.* As stressed in the section above on multiple equilibria, however, the logic here is different from a standard Diamond and Dybvig model, since in our environment the investors in a given weak bank would choose to run if and only if they expect the investors in the other weak banks to run.

Discussion. A number of recent legislative changes aim to promote financial stability by endowing financial intermediaries with increased contractual flexibility, which would allow them to react as soon as they start to experience distress. For example, "gates" and withdrawal fees in money market mutual funds, swing pricing in the mutual fund industry more generally, and the new bail-in rules in the US, Europe and elsewhere can all be interpreted as giving intermediaries the opportunity – but not necessarily the obligation – of imposing losses on all (or subset) of their investors if this is deemed desirable for the long term health of the institution. The hope of these legislative reforms is that these new "bail-in options" would not only be effective in mitigating fragility (or even preventing runs entirely), but in addition, would eliminate the need for taxpayers to finance a bailout or at least drastically reduce the cost of government's interventions.

For instance, one important purpose of the recent reforms to money market mutual funds in the U.S. is to reduce investors' incentive to redeem quickly and ahead of others (i.e. to run) when the fund is in distress. The imposition of fees and gates must be approved by a fund's board of directors, who are to use these tools only if this is determined to be in the best interest of their shareholders. Notice that from the perspective of our model, withdrawal fees and "gates" can be captured as setting lower payments in weak banks. Our results suggest that, in an environment characterized by limited commitment, asymmetric information and bailouts, these bail-in options may not be used an thus be ineffective in promoting financial stability. In the next section, we examine ways a policy maker might reduce the inefficiencies that arise in the competitive equilibrium in our model and promote financial stability.

5 Macroprudential policy

Since the equilibrium allocation studied in section 4 is always constrained inefficient and, in addition, may involve a welfare-decreasing run by investors on weak banks, it is natural to ask what types of prudential policy would be useful in this environment. In this section we examine such *macroprudential* policies. In particular *(i)* system-wide bail-in, *(ii)* selective bail-in, *(iii)* increasing the tax rate and *(iv)* eliminating bailouts. While none of these policies leads to the constrained efficient allocation derived in Section 3, each is capable of raising welfare in some situations.

5.1 Restring early payments

A government *regulator* can restrict the set of payments each banks is allowed to contract upon with its investors. In the discussion to follow, we focus on the bad aggregate state since this is the only state where the regulator will potentially choose to restrict bank payments. So suppose that when the aggregate state is bad the regulator dictates that each bank must set a payment that belongs to the set $\mathrm{C_R} \equiv \{c_1, c_2, ..., c_n\}$. The regulator is able to observe payments and enforce that these payment belong to $\mathrm{C_R}$. At the same time, this regulator **does not** observe bank-specific states and the rest of the model is the same as before. Namely, after the first θ withdrawals all bank-specific states become publicly known and the *fiscal authority* distributes the existing tax revenue between weak banks and the public good in order to attain the ex-post optimal distribution of resources. Importantly, the regulator is assumed to be in charge only during the first θ withdrawals, which is the time it takes the fiscal authority to learn the bank-specific states and begin resolution. The fiscal authority lacks commitment (just as before) and is not bound to follow policies that were instituted by the regulator.

A regulatory policy is a choice of $\mathrm{C_R}$. If the economy has multiple equilibria for given regulatory policy we assume that the equilibrium with a bank run is selected; the results would be qualitatively unchanged if we used some other equilibrium selection rule. Moreover, and without loss of generality, we restrict attention to sets $\mathrm{C_R}$ consisting of three elements 0, c_{1W} and c_{1S}. Thus, for each combination of σ_k and ρ_k bank-k can either suspend payments $c_1^k = 0$ or choose between c_{1W} and c_{1S}. Each bank will choose a contract that maximizes the sum of its investors expected utilities subject to regulatory constraints. In the bad aggregate state, the regulator will choose c_{1W} and c_{1S} in order to maximize the sum of all investors utilities anticipating the behavior of banks and the intervention of the fiscal authority. The following terminology will be useful for the analysis to follow.

Definition 2. The regulatory policy:

(i) Allows for a *full bail-in* if $c_{1W} = \hat{c}_1$ and a *partial bail-in* if $c_{1W} > \hat{c}_1$.

(ii) Triggers a *system-wide bail-in* if $c_{1W} = c_{1S} = \bar{c}$ and $\bar{c} < c_{1S}^*$.

(iii) Triggers a *selective bail-in* if $c_{1W} < c_{1S}$ and weak banks choose c_{1W} and sound banks choose c_{1S}.

Before proceeding we point out that the labels *regulator* and *fiscal authority* are just for concreteness. What is important is that the policy maker moves twice: first in its capacity as a regulator and later on in its capacity as a fiscal authority. The action taken as a fiscal authority, however, is chosen ex-post as a best respond to the situation at hand. Also, observe that the regulator was implicitly present in section 4 where we maintained that banks were not allowed to make a payment exceeding an upper bound $\bar{c}$. Moreover, this $\bar{c}$ was set to equal to c_{1S}^*, the payment made by sound banks in the constraint efficient allocation. Thus, the regulatory policy in section 4 can be viewed as a special case of the one in this section.

5.2 Bail-ins

Ideally, the regulatory policy would fully bail-in all weak banks while leaving all sound banks unaffected. Such a policy, however, cannot be implemented. The reason is that instead of fully bailing-in its investors, each weak bank strictly prefers to behave as a sound in expectation of a future bailout. Thus, the regulator has two options: impose a system-wide bail-in or ensure that the weak banks self-select into a bail-in. In order to induce weak banks to bail-in their investors the regulator commits to allow partial bail-ins. A partial bail-in is made attractive to weak banks by: *(i)* setting c_{1W} sufficiently low so that that a weak bank choosing to bail-in would prevent a run from its investors and at the same time *(ii)* setting c_{1W} sufficiently high so that the difference between c_{1S} and c_{1W} is not large enough to induce a given weak bank to incur the cost of run entailed by deviating to c_{1S}. The optimal regulatory policy is characterized next.

Proposition 12. *For some parameters, the regulator ensures that weak banks self-select into a partial bail-in. For other parameters, the regulator imposes a systemic bail-in. In some cases, the regulatory policy has a prudential impact by eliminating the run equilibrium.*

Proposition 12 is established in Figure 3 showing the optimal regulatory policy as a function of the size of the real losses per bank $\bar{\sigma}$ and the fraction of the banks experiencing real losses when the aggregate state is bad n. The remaining parameters are as follows $R = 1.5$, $\pi = \theta = 0.5$, $\delta = 0.5$, $\gamma = 6$, $q = 0.05$. Figure 3 shows that a system-wide bail-in dominates for some parameters, while a selective bail-in which ensures that weak banks self-select into a bail-in dominates for others. In particular, the dark region shows combinations of $\bar{\sigma}$ and n where it is optimal for the regulator to trigger a system-wide bail-in. The white region shows combinations of $\bar{\sigma}$ and n where the optimal regulatory policy ensures that weak banks self select into a partial bail-in.

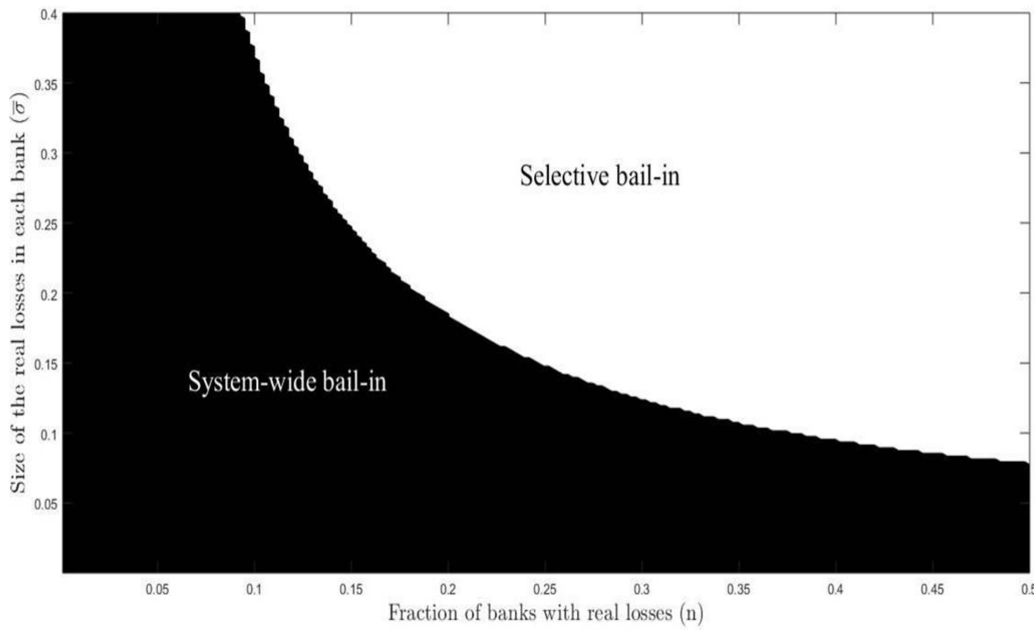


Figure 3: Optimal regulation

Prudential effect of bail-ins. For all combinations of n and $\bar{\sigma}$ shown on figure 3, the equilibrium is unique and such that weak banks do not experience a run. In order to establish that the regulator can have a prudential effect we need to show that the regulatory intervention eliminates the existence of the run equilibrium for some parameters. This can readily be done by referring back to Figure 2, which was constructed under the same parameters, but without the regulatory policy of this section. We can see that the regulatory policy eliminates the run equilibrium for combinations of n and $\bar{\sigma}$ belonging either to the gray region or to the white region of Figure 2.

Recall that a run in this environment is the result of an *externality* in payoffs across weak banks. That is, while it is optimal for weak bank k to keep its payoffs as high as possible, if all other weak banks behave in this way, then bank k ends up receiving a lower bailout payment, which in turn leads to a lower consumption allocation for the remaining investors in bank k. If this payoff externality is sufficiently severe, patient investors will choose to run on the bank in an attempt to withdraw before the government's intervention. One way to reducing this payoff externality is to trigger a system-wide bail-in (as in the black region of the figure), another is to ensure that weak banks self-select into a bail-in (as in the white region of the figure).

System-wide bail-in. For combinations of n and $\bar{\sigma}$ in the black region of Figure 3, the optimal regulatory policy is to trigger a system-wide bail-in by requiring all banks to set $\bar{c} < c_{1S}^*$. That is, the regulator imposes a cap on period payments $\bar{c}$ which applies equally to all banks. Recall that the inefficiency in our environment arises because of the (socially) excessive levels of maturity transformation undertaken by weak banks before being bailed out by the government. By imposing that all banks select the same payment which is below c_{1S}^* the government will bring their levels of maturity transformation closer to the socially desirable level. At the same time, a binding cap would also impose a cost on the sound banks since they will be prevented from setting c_{1S}^* in period 1. Proposition 12 shows that it can nevertheless be optimal for the government to impose a binding cap $\bar{c} < c_{1S}^*$ (i.e. a bail-in for all banks) since the negative effect on the sound banks will be more than offset by the lower misallocation of resources associated with the weak banks. In addition, the government would choose an even lower $\bar{c}$ if the misallocation resources stemming from the behavior of the weak banks is larger. This misallocation will be proportional to the fraction of weak banks n, the size of the losses in each weak bank $\bar{\sigma}$, the fraction of withdrawals which must take place before the government learns who are the weak banks θ.

Observe that in this environment a system-wide bail-in can also be interpreted as imposing restrictions on the dividends paid out by all banks during a period of financial stress. Alternatively, one can think of $\bar{c}$ as a contingent equity with a systemic trigger – if the aggregate state is bad (the systemic event) then all banks must bail-in their investors, even though the government realizes that not all banks are weak. Notice that the assumption that $\bar{c}$ is fixed before withdrawals begin is important. If the government were unable to commit to $\bar{c}$ and could change its value after withdrawal decisions have been made, a time-inconsistency problem would arise as the government would often prefer to set a higher cap. Hence, our model indicates that finding a way to commit to $\bar{c}$ by, for example, passing regulations which embed systemic triggers or similar measures is important for financial stability.[15]

Selective bail-in. For combinations of n and $\bar{\sigma}$ in the white region of Figure 3, the optimal regulatory policy is to ensure that weak banks self-select into a bail-in. A bank that voluntarily chooses to bail-in its investors (i.e. to set c_{1W} rather than the higher c_{1S}) must have an incentive to do so. A given weak bank will not self-select into a partial bail-in, unless the following two conditions are satisfied. First, bailing-in the first θ investors from c_{1S} to c_{1W} must prevent a run, that is, $c_{1W} \leq \hat{c}_2 < c_{1S}$. If that were not the case then a weak bank strictly prefers to set the higher c_{1S} (given that it always receives $(\hat{c}_1, \hat{c}_2)$ in resolution). Sec-

[15] This logic is similar to that in Ennis and Keister (2009), which shows how the inability to commit to suspend payments can create financial fragility.

ond, the difference between c_{1S} and c_{1W} cannot be too large, otherwise a weak bank would still prefer to set c_{1S}, even at the cost of a run. Formally, a weak bank would voluntarily bail-in to c_{1W} if the following condition is satisfied.

$$\theta u(c_{1W}) + (\pi - \theta)u(\hat{c}_1) + (1 - \pi)u(\hat{c}_2) \geq \theta u(c_{1S}) + (1 - \theta)\left[\pi u(\hat{c}_1) + (1 - \pi)u(\hat{c}_2)\right]$$

The bank can set the payment in period 1 to c_{1W} (i.e. bail-in), which would prevent a run and ensure that the sum of utilities for all investors in the bank equals the left hand side of the above expression. Alternatively, if the bank sets c_{1S}, then this would generate a run $c_{1S} > \hat{c}_2$ and deliver sum of utilities for its investors equal to the right hand side of the expression above. Notice that the above condition cannot be satisfied unless $c_{1W} > \hat{c}_1$. That is, a full bail-in is never privately optimal for a weak bank, even if this would prevent a run. This is a restatement of Proposition 8 which showed that the constrained efficient allocation - which required that weak banks enact a full bail-in - cannot be sustained in equilibrium. Also, observe that providing incentives for a weak bank to choose a (partial) bail-in can be attained either through lower c_{1S} or a higher c_{1W}. Operating on either margin will be costly for the regulator: lowering c_{1S} further below c_{1S}^* would increase the miss-allocation of resources in sound banks, at the same time, increasing c_{1W} (and thus decreasing the size of the bail-in) would further distort the allocation in weak banks.

In some cases, implementing a selective bail-in requires the regulator to operate on both margins. That is, allowing for a partial bail-in while distorting the allocation in sound banks in order to relax weak banks' incentive to behave as sound. Observe how the ability to commit ex-ante to the regulatory policy plays a crucial role in ensuring that weak banks voluntarily choose to bail-in their investors at the start of period 1. In particular, if the regulator allows for a partial bail-in, then weak banks will reveal themselves to the regulator by choosing c_{1W} rather than c_{1S}. At this point, the ex-post optimal action for the regulator would be to step in and mandate that each weak bank bails-in all the way to $\hat{c}_1$. Weak banks would anticipate this behavior and will set c_{1S} in order not to reveal themselves to the regulator. Therefore, a regulator that lacks commitment would fail to implement a selective bail-in and as a result will have to rely solely on a system-wide bail-in.

5.3 Increasing the tax rate

Previously the tax rate τ was set to its level in the constrained efficient allocation, τ^*. We relax this assumption in the current section and allow the amount of taxes collected in period 0 to serve as a macroprudential tool. Specifically, in period 0, the government chooses τ in order to maximize the sum of investors expected utilities. We have the following result.

Proposition 13. *If $q > 0$, the optimal choice of the tax rate in the decentralized economy is above its constrained efficient counterpart. That is, $\tau^D > \tau^*$. For some parameters, the selected tax rate τ^D eliminates the run equilibrium.*

Compared to the constrained efficient allocation, weak banks in the decentralized economy always choose to keep their early payments at an exceedingly high level (from a social perspective) and therefore will have fewer remaining resources available when bailout payments are made. This fact, in turn, would end up placing greater strain on the government whose objective is to use the available tax revenues τ both to support private consumption through the bailout transfers to weak banks and to provide the public good. Collecting more taxes in period 0 is one way to relax the budget constraint of the government in the bad aggregate state. Furthermore, the government has an added *prudential* motive to collect more taxes in period 0. In particular, a larger tax rate τ implies that the government would have more resources and provide larger bailouts to weak banks. Patient investors, in turn, anticipate that the bailout transfers will be more generous and - provided that these transfers are sufficiently high – will choose not to run on the weak banks. Thus, increasing τ would eventually have a prudential effect by eliminating the run equilibrium. At the same time, the opportunity cost of collecting more taxes is that fewer resources will be placed in the more productive private technology. Hence the government's willingness to increase the tax rate above its constrained efficient level depends on the ex-ante probability of the bad aggregate state q. The larger q is the more willing is the government to increase the tax rate above its constrained efficient level.

5.4 Eliminating bailouts

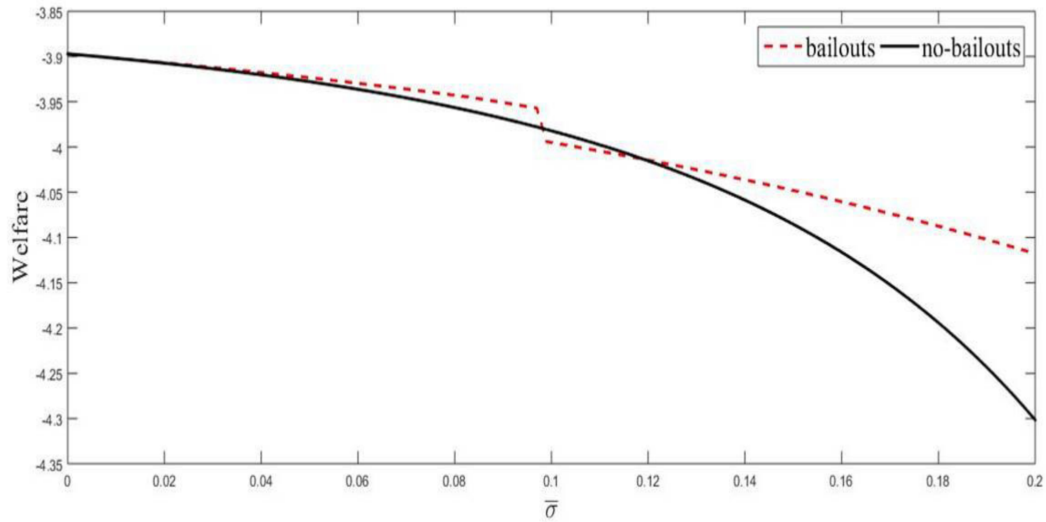
Suppose that in period 0 the government can commit to a strict *no-bailouts* policy. That is, there will be no bailouts, even after the government discovers the bank-specific states. Note that such a rule requires commitment, since ex-post the government would find it optimal to bail out weak banks not only because they might experience a run, but also because they have sustained real losses. The entire discussion in this section is therefore predicated on the ability of the government to pre-commit not to bail out weak banks. We can show that a no-bailout policy will prevent bank runs, but in many cases will lower welfare.

Proposition 14. *If the government can credibly commit not to bail out weak banks, then bank runs will not occur in equilibrium. For some parameters imposing a no-bailout rule would raise welfare, while for other parameters welfare would decrease.*

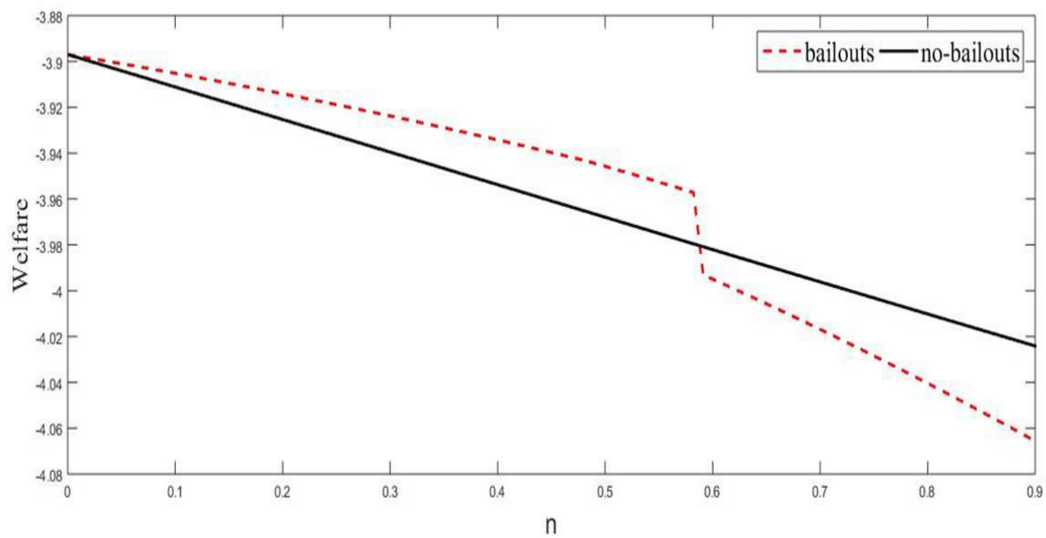
According to Proposition 14, the first implication of a (credible) no-bailout policy is that it will promote financial stability since runs can no longer occur as part of the equilibrium in the decentralized economy. The reason is that in our environment weak banks have all the information and contractual flexibility necessary to prevent runs and, in the absence of bailouts, will have no incentive to delay fully bailing-in their investors as soon as they have sustained real losses or realize that a run is underway. Proposition 14 also establishes that a no-bailout rule might improve ex-ante welfare for some parameter values. However, for other parameters eliminating bailouts would, in fact, result in inferior outcomes. The reason is that ex-post bailouts deliver a socially valuable transfer of tax revenue from the public sector to those banks that had sustained losses and thus ensuring a better allocation of resources between the investors in the weak banks and the public good. Given that banks can experience runs in addition to sustaining fundamental losses, such an ex-post co-insurance mechanism between the private and the public sector can be socially valuable even if it leads to moral hazard (as it invariably does in our environment).

Proposition 14 is demonstrated in panel (a) and panel (b) in Figure 4, where the y-axis in each panel plots the ex-ante welfare associated with the bailouts economy (the dashed red line) and with the no-bailouts economy (the black line). Focus first on panel (a) where the x-axis plots the real losses $\bar{\sigma}$ sustained by weak banks in the bad aggregate state.[16] We can see from the figure that the two economies deliver almost identical welfare for relatively small values of $\bar{\sigma}$. However, as $\bar{\sigma}$ is increasing beyond approximately 0.04 the bailouts economy starts to perform better until it becomes fragile at around 0.1 (hence the kink in the dashed red line). However, for sufficiently high levels of the real losses, the positive effect of ex-post bailouts (in terms of the improved allocation of resources between the public and the private good in the bad aggregate state) is relatively large. Thus permitting bailouts delivers higher ex-ante welfare, even if weak banks experience a run in the bailouts economy. Next, focus on panel (b) on the figure where the x-axis plots the fraction of banks suffering real losses, n. We can see that for sufficiently high values of n prohibiting bailouts would *both* promote stability and increase ex-ante welfare. The reason for this result is the following: holding fixed the level of real losses $\bar{\sigma}$, the benefit of promoting stability will be high if the fraction of banks the can experience a run n is high. As panel (b) on figure 4 shows, in such cases, it can it be desirable to pre-commit to a strict no-bailouts policy, even at the cost of forsaking the socially valuable ex-post function of the government's bailouts.

[16] The remaining parameters for panel (a) in Figure 4 are $R = 1,5$, $\pi = 0.5$, $\gamma = 5$, $\delta = 0.5$, $\theta = 0.5$, $q = 0.05$ and $n = 0.6$. The remaining parameters for panel (b) are $R = 1,5$, $\pi = 0.5$, $\gamma = 5$, $\delta = 0.5$, $\theta = 0.5$, $q = 0.05$ and $\bar{\sigma} = 0.1$.



(a)



(b)

Figure 4: Welfare in the no-bailouts economy.

Summarizing. The main insight from this section is that policies to promote better outcomes in this environment must take into account banks' incentives. A bank that anticipates to be bailed out has an incentive **not to** bail-in its investors after experiencing losses. One way for the government to address this issue is to trigger a *system wide bail-in* whenever a fraction of the banks are in distress. Such a policy leaves weak banks with no option but to bail-in their investors, but it also comes at the cost. Namely, sound banks would also be forced to enact a bail-in which will be unnecessary in their case. A different policy intervention is to implement a *selective bail-in.* The challenge in this case is to induce weak banks to bail-in as soon as they experience losses. The government can accomplish such a goal by allowing for a partial bail-in and, in some cases, by distorting the allocation in the sound banks (i.e. by mandating that they set a payment different than c^*_{1S}). Proposition 12 showed that a selective bail-in will dominate for some parameters whereas a system-wide bail-in will dominate for other. Another way to increase welfare in this environment is to collect more taxes in period 0 by *increasing the tax rate* beyond its level in the constrained efficient allocation. Higher tax revenue can have a prudential impact by relaxing the budget constraint of the government. A final option is to *prohibit bailouts.* A no-bailout policy ensures that runs do not occur but would also dispense with the socially valuable ex-post insurance function of the government's bailouts intervention. The effect on ex-ante welfare is thus ambiguous and depends on parameter values. In addition, a no-bailout policy might be plagued by government credibility issues. None of the above policies would restore the constraint efficient allocation, but each one of them is capable of preventing runs and raising welfare under certain conditions.

6 Conclusion

A necessary ingredient for a bank run to occur in the the bank in question be slow to react to the surge in withdrawal demand. This slow reaction is what leads investors to anticipate that the future payments made by the bank will be smaller and, hence, gives them an incentive to try to withdraw before the reaction comes. In the previous literature, the primary factors behind this slow response have been exogenously imposed rather than derived endogenously as part of the equilibrium outcome. Specifically, banks' failure to respond in a timely manner has been justified by assuming that either (i) contracts are rigid and therefore cannot be ex-post altered to deal with a run, or (ii) banks were unable to respond efficiently to a run because they were (at least initially) unaware that the run was actually taking place.

In contrast, we have presented a model of banking and government interventions where (i) banks maximize the utilities of their investors (i.e. there are no agency costs), (ii) contracts can be made fully state contingent, and (iii) banks always have sufficient information to respond in a timely and effective way to an incipient run. In common with the existing literature, a bank run in our setting can occur only when the bank's reaction to the run is delayed. However, the delayed reaction in our model is the endogenous choice of the bank, acting in the best interests of its investors. We show that this framework has a number of interesting implications. For example, banks will not have an incentive to use bail-in options and similar measures to impose discretionary losses on their investors when they anticipate to be bailed out by the government later on. As a result, the new bail-in rules might turn out to be not as effective in promoting financial stability as they were originally expected to be. We addressed some of the possible approaches to fix these weaknesses of the "bail-in rules" and concluded that they are unlikely to solve the problem of bank runs on their own due to a combination of asymmetric information and the policy maker's lack commitment.

References

[1] Allen, F. and Gale, D. (2004) "Financial fragility, liquidity, and asset prices" *Journal of the European Economic Association,* **2**, 1015-1048.

[2] Andolfatto, D., Nosal, E., and Sultanum, B. (2016) "Preventing bank runs," *Theoretical Economics*, forthcoming.

[3] Bianchi, J. (2012), "Efficient bailouts?" NBER Working Paper No. 18587.

[4] Boyd, J.H., Chang, C. and Smith, B.D. (2002) "Deposit insurance: A reconsideration" *Journal of Monetary Economics,* **49,** 1235-1260.

[5] Brown, C. and Dinç, S, (2005), The politics of bank failures: Evidence from emerging markets, *The Quarterly Journal of Economics*, 120, issue 4, p. 1413-1444

[6] Calomiris, C.W., and Kahn, C.M. (1991) "The role of demandable debt in structuring optimal banking arrangements" *American Economic Review,* **81**, 497-513.

[7] Chari, V.V. and Kehoe, P.J. (2016) "Bailouts, time inconsistency, and optimal regulation: A macroeconomic view" *American Economic Review,* **106**, 2458-93.

[8] Cooper, R. and Kempf, H. (2013) "Deposit insurance and bank liquidation without commitment: Can we sleep well?" NBER Working Paper No. 19132.

[9] Diamond, D.W. and Dybvig, P.H. (1983) "Bank runs, deposit insurance, and liquidity" *Journal of Political Economy,* **91**, 401-419.

[10] Diamond, D.W. and Rajan, R.G. (2001) "Liquidity risk, liquidity creation and financial fragility: A theory of banking" *Journal of Political Economy*, **109**, 287-327.

[11] Ennis, H.M. (2012) "Some theoretical considerations regarding net asset values for money market funds" Federal Reserve Bank of Richmond *Quarterly Review*, **98**, 231-254.

[12] Ennis, H.M. and Keister, T. (2009) "Bank runs and institutions: The perils of intervention" *American Economic Review,* **99**, 1588-1607.

[13] Ennis, H.M. and Keister, T. (2010) "Banking panics and policy responses" *Journal of Monetary Economics,* **57**, 404-419.

[14] Farhi, E. and Tirole, J. (2012) "Collective moral hazard, maturity mismatch and systemic bailouts" *American Economic Review,* **102**, 60-93.

[15] Goldstein, I. and Pauzner, A. (2005) "Demand deposit contracts and the probability of bank runs" *Journal of Finance,* **60**, 1293-1327.

[16] Green, E.J. (2010) "Bailouts" Federal Reserve Bank of Richmond *Economic Quarterly,* **96**,11-32.

[17] Green, E.J. and Lin, P. (2003) "Implementing efficient allocations in a model of financial intermediation" *Journal of Economic Theory,* **109**, 1-23.

[18] Keister T. (2016) "Bailouts and Financial Fragility" *Review of Economic Studies*, **83,** 704–736.

[19] Kroszner, R. and Strahan, P. 1996. "Regulatory incentives and the thrift crisis: Dividends, mutual-to-stock conversions, and financial distress." Journal of Finance **51**, 1285–320.

[20] Martin, A. (2006) "Liquidity provision vs. deposit insurance: preventing bank panics without moral hazard" *Economic Theory,* **28**, 197–211.

[21] Peck, J. and Shell, K. (2003) "Equilibrium bank runs" *Journal of Political Economy,* **111**, 103-123.

[22] Rogoff, K. and Sibert, A (1988) "Elections and macroeconomic policy cycles" *Review of Economic Studies* **55**, Jan. 1988, 1-16

[23] Wallace, N. (1988) "Another attempt to explain an illiquid banking system: the Diamond and Dybvig model with sequential service taken seriously" Minneapolis Fed *Quarterly Review,* **12** (Fall), 3-16.

[24] Wallace, N. (1990) "A banking model in which partial suspension is best" Minneapolis Fed *Quarterly Review,* **14** (Fall), 11-23.

[25] Walther, A. and White, L. (2017) "Optimal bank resolution" mimeo, April.

Appendix: Proofs

Proposition 1.

Proof. We want to show that sound banks do not receive a bailout and the consumption of investors in sound banks is independent of the aggregate stat. In order to establish the desired result, to it suffices to show that $b_S^* = 0$, since the following condition can be shown to hold:

$$b_S^* = 0 \quad \Leftrightarrow \quad (c_{10}^*, c_{20}^*) = (c_{1S}^*, c_{2S}^*) \tag{27}$$

The above follows from the feasibility constraints in (2) and (3) and the first order conditions in (6). Next, in order to establish that $b_S^* = 0$ we proceed by contradiction: suppose that sound banks receive a bailout in the bad aggregate state, that is, $b_S > 0$. From (8) and the feasibility constraints we obtain that weak banks must also receive a bailout, $b_W > 0$. Moreover, b_S^* and b_W will be set to ensure that the marginal value of the resources on all banks with a bailout equals the marginal value of the public good:

$$\mu_S = \mu_W = v'(\tau - (1-n)b_S - nb_W)$$

From the above, the condition for the optimal choice of the tax rate τ in (7) reduces to:

$$v'(\tau) = \mu_S$$

However, we also have $(1-n)b_S + nb_W > 0$, that is, the total bailout in the bad aggregate state is strictly positive and therefore:

$$v'(\tau) < v'(\tau - (1-n)b_S - nb_W)$$

Hence, we obtain $\mu_S < \mu_S$ - a contradiction. Therefore we must have $b_S^* = 0$ i.e. sound banks are not bailed out even in the bad aggregate state. Finally, from (27) we obtain that $(c_{10}^*, c_{20}^*) = (c_{1S}^*, c_{2S}^*)$, namely the payment profile in sound banks is independent of the aggregate state. □

Proposition 2.

Proof. Before establishing the desired result, it will be useful for future reference to collect all conditions characterizing the constrained efficient allocation. First, sound banks do not receive a bailout and provide identical allocation in each of the two aggregate states

(Proposition 1). This allocation, denoted c_{1S}^* and c_{2S}^*, is characterized by:

$$\pi c_{1S}^* + (1-\pi)\frac{c_{2S}^*}{R} = 1 - \tau^* \quad \text{and} \quad u'\left(c_{1S}^*\right) = Ru'\left(c_{2S}^*\right) \tag{28}$$

Second, weak banks are bailed out by the government and their allocation, denoted c_{1W}^* and c_{2W}^*, will be characterized by:

$$\pi c_{1W}^* + (1-\pi)\frac{c_{2W}^*}{R} = 1 - \tau^* - \bar{\sigma} + b^* \quad \text{and} \quad u'\left(c_{1W}^*\right) = Ru'\left(c_{2W}^*\right) \tag{29}$$

where b^* is the per-investor bailout in each weak bank. Third, the per-investor bailout in weak banks b^* and the optimal choice of the tax rate τ^* is characterized by:

$$\mu_W^* = v'\left(\tau^* - nb^*\right) \quad \text{and} \quad v'\left(\tau^*\right) = u'\left(c_{1S}^*\right) - \frac{q(1-n)}{1-q}\left(u'\left(c_{1W}^*\right) - u'\left(c_{1S}^*\right)\right) \tag{30}$$

The constrained efficient allocation is thus the solution to (28) - (29) and can be summarized in the following vector:

$$\text{CE} \equiv \left(c_{1S}^*, c_{2S}^*, c_{1W}^*, c_{2W}^*, b^*, \tau^*\right)$$

Finally, let μ_S^* and μ_W^* denote the shadow value of the resources in sound and weak banks respectively. That is,

$$\mu_S^* = u'\left(c_{1S}^*\right) \quad \text{and} \quad \mu_W^* = u'\left(c_{1W}^*\right)$$

We want to show $(c_{1S}^*, c_{2S}^*) \gg (c_{1W}^*, c_{2W}^*)$, suppose this were not the case, i.e. assume that

$$c_{1S}^* \leq c_{1W}^* \quad \text{and} \quad c_{2S}^* \leq c_{2W}^* \tag{31}$$

Since $\bar{\sigma} > 0$, from (31) and (29) it follows that weak banks must be bailed out $b^* > 0$. From proposition 1 the first order condition for the optimal choice of τ can now be written as:

$$v'(\tau^*) = \mu_S^* - \frac{q(1-n)}{1-q}\left(\mu_W^* - \mu_S^*\right)$$

In addition, if (31) is true then we must have $\mu_W^* \leq \mu_S^*$. Moreover the following must also hold:

$$
\begin{aligned}
\mu_W^* &= v'\left(\tau^* - nb^*\right) \\
&> v'(\tau^*) \\
&= \mu_S^* - \frac{q(1-n)}{1-q}\left(\mu_W^* - \mu_S^*\right) \\
&\geq \mu_S^* \\
&\geq \mu_W^*
\end{aligned}
$$

That is, we obtain $\mu_W^* > \mu_S^*$ which contradicts $\mu_W^* \leq \mu_S^*$. Therefore (31) cannot be true and we must have

$$
(c_{1S}^*, c_{2S}^*) \gg (c_{1W}^*, c_{2W}^*) \quad \text{and} \quad \mu_W^* > \mu_S^*
$$

which establishes the first part of the proposition. Next, we want to show that $b^* > 0$. So, let $\mu_W^*(b)$ be the shadow value of the resources in a given weak bank resulting from a per-investor bailout of b. The following must be satisfied:

$$
\mu_W(b^*) > v'(\tau^*) > \mu_S^*
$$

At the same time, if $b^* = 0$, then it must be the case that:

$$
v'(\tau^*) \geq \mu_W\left(b^*\right) \quad \text{for} \quad b^* = 0
$$

Otherwise, each weak bank will receive a strictly positive bailout. Hence, if $b^* = 0$ it has to be the case that $v'(\tau^*) \geq \mu_W\left(b^*\right)$ and $v'(\tau^*) < \mu_W\left(b^*\right)$ which is a contradiction since both of this conditions cannot hold at the same time. Therefore, our initial assumption that $b^* = 0$ cannot be true and we must have $b^* > 0$. □

Proposition 3.

Proof. We must show that for every allocation $c_S = (\rho_W,\, c_{1S},\, c_{2S})$ for which both (10) and (11) is true, we can find a banking contract that would lead to this allocation as the *unique equilibrium* of the withdrawal game for the investors in the bank, conditional on the bank being sound. So suppose that

$$
\pi c_{1S} + (1-\pi)\frac{c_{2S}}{R} = 1 - \tau
$$

$$
\rho_S = \pi \quad \text{and} \quad c_{1S} \leq c_{2S}
$$

Then given the following contract, the only equilibrium when the bank is sound is for all its investors to follow the no-run strategy.

$$c_1^k\left(0,\,\rho^k\right)=\left\{\begin{array}{c}c_{1S}\\0\end{array}\right\}\quad\text{if}\quad\left\{\begin{array}{c}\rho^k=\pi\\\rho^k>\pi\end{array}\right\}$$

If all investors follow the no-run strategy (i.e. $x_0^k=0$) then period-1 requests for withdrawals equal $\rho^k=x_0^k+(1-x_0^k)\pi$ and the contract specifies $c_1^k\left(0,\,\rho^k\right)=c_{1S}$, whereas the budget constraint implies $c_2^k\left(0,\,\rho^k\right)=c_{2S}$, where c_{2S} is obtained from the bank's budget constraint. Since $c_{1S}\leq c_{2S}$ all investors are best responding with the no-run strategy. Specifically, if patient, an investor does not gain from withdrawing in period 1 and therefore best responds by withdrawing in period 2 as specified in the no-run strategy.

On the other hand, for any x_0^k such that $0<x_0^k\leq 1$, that is, if positive measure of the investors in the bank follow the run strategy in case the bank is sound, then $\rho^k>\pi$ and the above contract would sets $c_1^k\left(0,\,\rho^k\right)=0$. In this case, the bank's budget constraint in (10) yields $c_1^k\left(0,\,\rho^k\right)<c_2^k\left(0,\,\rho^k\right)$. That is, a positive measure $x_0^k(1-\pi)>0$ of the investors (i.e. those that are patient *and* follow the run strategy) are not best responding. Following an argument analogous to the one above we can establish that a banking contract with the property:

$$c_1^k\left(0,\,\rho^k\right)=\left\{\begin{array}{c}B\\c_{1S}\\0\end{array}\right\}\quad\text{if}\quad\left\{\begin{array}{c}\rho^k<\rho_S\\\rho^k=\rho_S\\\rho^k>\rho_S\end{array}\right\}$$

where $B>c_{2S}$, leads to an allocation satisfying (10), and either $\rho_S\in[\pi,1]$ and $c_{1S}=c_{2S}$ or $\rho_S=1$ and $c_{1S}\geq c_{2S}$ as the unique equilibrium of the withdrawal game of the bank's investors when the bank's fundamental is sound. □

Lemma 1.

Proof. If ρ_S is equal to π then the solution to the problem of maximizing (12) subject to (10) is given by:

$$u'\left(c_{1S}^*\right)=Ru'\left(c_{2S}^*\right)\quad\text{and}\quad\pi c_{1S}^*+(1-\pi)\frac{c_{2S}^*}{R}=1-\tau \tag{32}$$

Since $R>1$ we have $c_{1S}^*<c_{2S}^*$. That is, the allocation $(\rho_S^*,c_{1S}^*,c_{2S}^*)$ characterized by (32) and $\rho_S^*=\pi$ satisfy both (10) and (11). Next, we show that allocation $(\rho_S^*,c_{1S}^*,c_{2S}^*)$, in fact, maximizes the program in (12). To see that, consider any other allocation (ρ_S,c_{1S},c_{2S}) such that $\rho_S=\pi$ and

$$c_{1S} \leq c_{2S} \quad \text{and} \quad \pi c_{1S} + (1-\pi)\frac{c_{2S}}{R} = 1-\tau$$

That is, the allocation satisfies (10) and (11) and therefore:

$$V_S\left(c^*_{1S},\, \pi\right) > V_S\left(c_{1S},\, \pi\right) \tag{33}$$

Next, define the function $\bar{c}(\rho_S)$:

$$\bar{c}(\rho_S) \equiv \frac{1-\tau}{\rho_S + (1-\rho_S)R^{-1}}$$

For each $\rho_S \in [\pi, 1)$ the allocation $(\rho_S,\, \bar{c}(\rho_S),\, \bar{c}(\rho_S))$ satisfies both (10) and (11) and thus is a potential candidate for the program in (12). However, since $\bar{c}(\rho_S)$ is a strictly decreasing function of ρ_S it follows that $V_S\left(\bar{c}(\rho_S),\, \rho_S\right)$ is strictly decreasing in ρ_S. That is,

$$V_S\left(\bar{c}(\pi),\, \pi\right) > V_S\left(\bar{c}(\rho_S),\, \rho_S\right) \tag{34}$$

Combining (33) and (34) yields the desired result, namely the allocation in (32) maximizes the function (12) subject to (10) and (11). □

Proposition 4.

Proof. Suppose that in equilibrium bank-k's contract is given by c_1^k and the fraction of the bank's investors following the run strategy is $x^k_{\sigma_k}$ for $\sigma_k \in \{0,\, \bar{\sigma}\}$. If the bank is sound, $\sigma_k = 0$, then investors' withdrawal strategies imply that the request for period-1 withdrawals will be equal to

$$\rho_0^k = x_0^k + (1-x_0^k)\pi \geq \pi \tag{35}$$

Given ρ_0^k, the bank's contract would specify a period-1 payment of $c_1^k\left(\rho_0^k,\, 0\right)$ and a period 2 payment of

$$c_2^k\left(\rho_0^k,\, 0\right) = \frac{R\left[1-\tau-\rho_0^k c_1^k\left(\rho_0^k,\, 0\right)\right]}{1-\rho_0^k}$$

The resulting allocation will be consistent with equilibrium if it also satisfies the implementability constraint in (11). That is

$$\rho_0^k = \left\{\begin{array}{c} \pi \\ \in [\pi, 1] \\ 1 \end{array}\right\} \quad \text{as} \quad c_1^k\left(\rho_0^k,\, 0\right) \left\{\begin{array}{c} \leq \\ = \\ \geq \end{array}\right\} c_2^k\left(\rho_0^k,\, 0\right) \tag{36}$$

The sum of utilities for the investors in the bank associated with the allocation $\left(\rho_0^k,\, c_1^k\left(\rho_0^k,\, 0\right),\, c_2^k\left(\rho_0^k,\, 0\right)\right)$ will be given by:

$$\rho_0^k u\left(c_1^k\left(\rho_0^k,\, 0\right)\right) + (1-\rho_0^k) u\left(c_2^k\left(\rho_0^k,\, 0\right)\right) \tag{37}$$

Applying Lemma 6, the expression in (37) subject to (35) - (36) is maximized only if the bank's allocation is characterized by $\rho_S^* = \pi$ and (32). The allocation in (32) will be implemented if the bank's contract is:

$$c_1^k\left(0,\, \rho^k\right) = \left\{ \begin{array}{c} c_{1S}^* \\ 0 \end{array} \right\} \quad \text{for} \quad \left\{ \begin{array}{c} \rho^k = \pi \\ \rho^k \neq \pi \end{array} \right\} \tag{38}$$

Indeed, above contract leads to the allocation in Lemma 6 as the *unique equilibrium* when the bank is sound. Hence we obtain the desired result, namely, sound banks do not experience a run $\rho_0^k = \pi$ (i.e. $x_0^k = 0$) and their consumption allocation (c_{1S}^*, c_{2S}^*) is the same as in the constrained efficient case. □

Proposition 5.

Proof. Consider an allocation $(\rho_W,\, c_{1W})$ which satisfies the implementability constraint in (22). We must consider three cases: (i) $\rho_W = \pi$ and $c_{1W} \leq \hat{c}_2$. (ii) $\pi < \rho_W < 1$ and $c_{1W} = \hat{c}_2$. (iii) $\rho_W = 1$ and $c_{1W} \geq \hat{c}_2$. If $(\rho_W,\, c_{1W})$ is in case (i), that is, $\rho_W = \pi$ and $c_{1W} \leq \hat{c}_2$ then consider the following contract for bank-k

$$c_1^k\left(\bar{\sigma},\, \rho^k\right) = \left\{ \begin{array}{c} c_{1W} \\ 0 \end{array} \right\} \quad \text{for} \quad \left\{ \begin{array}{c} \rho^k = \pi \\ \rho^k > \pi \end{array} \right\} \tag{39}$$

If $x_{\bar{\sigma}}^k = 0$, i.e. if all investors in bank-k follow the no-run strategy, then $\rho^k = \pi$ and $c_1^k\left(\bar{\sigma},\, \rho^k\right) = c_{1W}$ and since $c_{1W} \leq \hat{c}_2$ it follows that all investors in bank-k are best responding with the no-run strategy.

On the other hand, for any $x_{\bar{\sigma}}^k > 0$, i.e. if the measure of investors in bank-k following the run strategy is positive, then $\rho^k = x_{\bar{\sigma}}^k + (1 - x_{\bar{\sigma}}^k)\pi > \pi$ and its contract specifies $c_1^k\left(\bar{\sigma},\, \rho^k\right) = 0 < \hat{c}_2$, which violates the the implementability constraint in (22). In particular, a fraction $x_{\bar{\sigma}}^k(1-\pi)$ of the investors in the bank (those that are both patent and follow the run strategy) will not be best responding.

Then, conditional on the bank being weak, an allocation such that $\rho_W = \pi$ and $c_{1W} \leq \hat{c}_2$ will be uniquely implemented by the contract in (39).

If $(\rho_W,\, c_{1W})$ is in case (ii), that is, $\pi < \rho_W < 1$ and $c_{1W} = \hat{c}_2$. Consider the following contract for bank-k

$$c_1^k\left(\bar{\sigma},\,\rho^k\right)=\left\{\begin{array}{c}\hat{c}_2+\epsilon\\ c_{1W}\\ 0\end{array}\right\}\quad\text{for}\quad\left\{\begin{array}{c}\rho^k<\rho_W\\ \rho^k=\rho_W\\ \rho^k>\rho_W\end{array}\right\}\tag{40}$$

If $x_{\bar{\sigma}}^k = z$, where z is such that $\rho^k = z + (1-z)\pi = \rho_W$, we have $c_1^k\left(\bar{\sigma},\,\rho^k\right) = c_{1W}$. Since $c_{1W} = \hat{c}_2$ all patient investors are indifferent between withdrawing in period 1 and period 2 and therefore they best respond with the run strategy.

On the other hand, for any $x_{\bar{\sigma}}^k \neq z$ we have $c_1^k\left(\bar{\sigma},\,\rho^k\right) = \hat{c}_2+\epsilon > \hat{c}_2$ for $\rho^k = z+(1-z)\pi < \rho_W$ and $c_1^k\left(\bar{\sigma},\,\rho^k\right) = 0 < \hat{c}_2$ for $\rho^k = z + (1-z)\pi > \rho_W$. In either case, $x_{\bar{\sigma}}^k \neq z$ will not be consistent with equilibrium.

Hence, we have shown that conditional on the bank being weak, an allocation such that$\pi < \rho_W < 1$ and $c_{1W} = \hat{c}_2$ will be uniquely implemented by the contract in (40).

Finally, if $(\rho_W,\, c_{1W})$ is in case (iii), that is, $\rho_W = 1$ and $c_{1W} \geq \hat{c}_2$. Consider the following contract for bank-k

$$c_1^k\left(\bar{\sigma},\,\rho^k\right)=\left\{\begin{array}{c}\hat{c}_2+\epsilon\\ c_{1W}\end{array}\right\}\quad\text{for}\quad\left\{\begin{array}{c}\rho^k<1\\ \rho^k=1\end{array}\right\}\tag{41}$$

In this case, we can readily verify that the only equilibrium when outcome k is weak will be for all its investors to follow the run strategy $x_{\bar{\sigma}}^k = 1$, which implies the allocation such that $\rho_W = 1$ and $c_{1W} \geq \hat{c}_2$ is uniquely implemented by the contract in (41). □

Proposition6.

Proof. Since $\hat{c}_1 < \hat{c}_2$ and $\dfrac{\partial\hat{\rho}_{\bar{\sigma}}}{\partial\rho_W} > 0$ it follows that

$$\frac{\partial V_W\left(c_{1W},\,\rho_W\right)}{\partial c_{1W}} > 0 \quad\text{and}\quad \frac{\partial V_W\left(c_{1W},\,\rho_W\right)}{\partial\rho_W} < 0 \tag{42}$$

First, if $c_{1W} < \hat{c}_2$, the implementability constraint (22) implies that $\rho_W = \pi$ and from (42)

$$V_W\left(c_{1W},\,\pi\right) < V_W\left(\hat{c}_2,\,\pi\right) \quad\text{for}\quad 0 < c_{1W} < \hat{c}_2$$

Second, if $c_{1W} = \hat{c}_2$ then from (22) we have $0 \leq \rho_W \leq 1$ and from (42)

$$V_W\left(\hat{c}_2,\,\rho_W\right) < V_W\left(\hat{c}_2,\,\pi\right) \quad\text{for}\quad \rho_W > \pi$$

Third, if $c_{1W} > \hat{c}_2$, then (22) implies that $\rho_W = 1$ and from (42)

$$V_W\left(c_{1W},\,1\right) < V_W\left(\bar{c},\,1\right) \quad\text{for}\quad \hat{c}_2 < c_{1W} < \bar{c}$$

where $\bar{c}$ is the maximum payment banks are allowed to make in period 1. The first two inequalities imply

$$V_W(c_{1W}, \rho_W) < V_W(\hat{c}_2, \pi) \quad \text{for} \quad c_{1W} \leq \hat{c}_2 \quad \text{and} \quad \rho_W > \pi$$

Therefore, if $\bar{c} < \hat{c}_2$, the bank will set $c_{1W} = \bar{c}$. On the other hand, if $\bar{c} \geq \hat{c}_2$, the bank sets:

(i) $c_{1W} = \bar{c}$ if $V_W(\bar{c}, 1) > V_W(\hat{c}_2, \pi)$.

(ii) $c_{1W} \in \{\bar{c}, \hat{c}_2\}$ if $V_W(\bar{c}, 1) = V_W(\hat{c}_2, \pi)$.

(iii) $c_{1W} = \hat{c}_2$ if $V_W(\bar{c}, 1) < V_W(\hat{c}_2, \pi)$.

Thus we obtain the desired result, namely c_{1W} will be set as high as possible or equal to $\hat{c}_2$. □

Proposition 7.

Proof. The payment in period 1 set by weak banks, c_{1W}, is is either equal to $\bar{c}$ or $\hat{c}_2$. If this were not the case, then Proposition 6 implies that banks' contracts are, in fact, not optimal. Also, the maximum payment banks are permitted to set in period-1 is capped at c_{1S}^*, which is the payment in period 1 set by the sound banks (that is, $\bar{c} = c_{1S}^*$). Hence, in equilibrium we have $c_{1W} \in \{c_{1S}^*, \hat{c}_2\}$. First, suppose that

$$c_{1S}^* > \hat{c}_2$$

and consider the following: (i) weak banks optimize by setting c_{1S}^*. Then since $c_{1S}^* > \hat{c}_2$ their investors best respond with the run strategy ($x_{\bar{\sigma}} = 1 \Rightarrow \rho_W = 1$). According to Proposition 6, weak banks would behave optimally if

$$V_W(c_{1S}^*, 1) \geq V_W(\hat{c}_2, \pi)$$

An optimal banking contract c_1^k in this case is given by (38) and (41), where $c_{1W} = c_{1S}^*$. Another possibility is (ii) weak banks optimize by setting $\hat{c}_2$. Then since $c_{1S}^* > \hat{c}_2$ their investors best respond with the no-run strategy ($x_{\bar{\sigma}} = 0 \Rightarrow \rho_W = \pi$). According to Proposition 6, weak banks would behave optimally if

$$V_W(c_{1S}^*, 1) \leq V_W(\hat{c}_2, \pi)$$

An optimal banking contract c_1^k in this case is given by (38) and (39), with $c_{1W} = \hat{c}_2$. On the other hand, suppose that

$$c_{1S}^* \leq \hat{c}_2$$

then Proposition 6 implies that weak banks optimize by setting $c_{1W} = c^*_{1S}$. Also, the fact that $V_W(c_{1W}, \rho_W)$ is decreasing in the second argument, yields:

$$V_W(c_{1W}, \pi) > V_W(c_{1W}, \rho_W)$$

for $\rho_W > \pi$ and therefore an allocation such that $\rho_W > \pi$ is not optimal for banks when $c^*_{1S} \leq \hat{c}_2$. An optimal banking contract c^k_1 in this case is given by (38) and (39), with $c_{1W} = c^*_{1S}$. □

Proposition 8.

Proof. A necessary condition for the equilibrium to be constrained efficient is for weak banks to pay the same amount to all investors withdrawing in period 1, $c_{1W} = \hat{c}_1$. That is, from the start of period 1, weak banks must lower their payments all the way down to their level in resolution $\hat{c}_1$. From Proposition 4, however, the equilibrium value of c_{1W} would equal either c^*_{1S} or $\hat{c}_2$. Then since $\min\{c^*_{1S}, \hat{c}_2\} > \hat{c}_1$ it follows that the equilibrium is not constrained efficient. □

Proposition 9.

Proof. The tax rate is fixed to its level in the constrained efficient allocation $\tau = \tau^*$ and the government imposed cap on early payments is set to $\bar{c} = c^*_{1S}$ - the early payment made by sound banks in the constrained efficient economy. First, we establish the existence of equilibrium runs in this environment.

I. *Equilibrium with a run.* Suppose that each bank k sets the following contract:

$$c^k_1\left(0, \rho^k\right) = \left\{ \begin{array}{c} c^*_{1S} \\ 0 \end{array} \right\} \quad \text{if} \quad \left\{ \begin{array}{c} \rho^k = \pi \\ \rho^k > \pi \end{array} \right\} \tag{43}$$

$$c^k_1\left(\bar{\sigma}, \rho^k\right) = \max\{c^*_{1S}, \hat{c}_2\} \quad \text{for} \quad \rho^k \in [\pi, 1] \tag{44}$$

Such a contract ensures that: (i) conditional on being sound, bank-k does not experience a run and provides the constrained efficient allocation (as must be necessarily the case from Proposition 4). (ii) Conditional on being weak, bank-k experiences a run if $c^*_{1S} > \hat{c}_2$. In order for this contract to be consistent with equilibrium the following condition must be satisfied (see Proposition 7):

$$u(\max\{c^*_{1S}, \hat{c}_2\}) - u(\hat{c}_2) \geq (1-\pi)(u(\hat{c}_2) - u(\hat{c}_1)) \tag{45}$$

This condition ensures that bank-k cannot deliver strictly higher welfare to its investors by deviating to a run proof contract. We can restrict attention to $c_{1S}^* > \hat{c}_2$, otherwise the the above weak inequality cannot be satisfied. The payments in resolution $(\hat{c}_1, \hat{c}_2)$ are determined from:

$$\theta c_{1S}^* + (\pi - \theta)\hat{c}_1 + (1-\pi)\frac{\hat{c}_2}{R} = 1 - \tau - \bar{\sigma} + \hat{b}$$

$$u'(\hat{c}_1) = Ru'(\hat{c}_2) = v'\left(\tau^* - n\hat{b}\right)$$

Given parameter values R, π, γ, δ, θ, q, n, $\bar{\sigma}$ and investor preferences as in (26) we solve for the equilibrium allocation resulting from banks choosing a contract as in (43) and (44) and then check if the condition in (45) holds. If this condition is satisfied, then the equilibrium where weak banks experience a run exists. The existence of such run equilibrium is numerically established on Figure 2 . Henceforth, we collect the allocation in the run equilibrium in the following vector:

$$\mathrm{D^R} \equiv \left(\tau^*,\, c_{1S}^*,\, c_{2S}^*,\, \hat{c}_1^R,\, \hat{c}_2^R,\, \hat{b}^R\right)$$

II. *Equilibrium without a run.* Suppose that the contract in all banks is the following:

$$c_1^k\left(0,\, \rho^k\right) = \left\{\begin{array}{c} c_{1S}^* \\ 0 \end{array}\right\} \quad \text{if} \quad \left\{\begin{array}{c} \rho^k = \pi \\ \rho^k > \pi \end{array}\right\} \tag{46}$$

$$c_1^k\left(\bar{\sigma},\, \rho^k\right) = \left\{\begin{array}{c} \min\{c_{1S}^*,\, \hat{c}_2\} \\ 0 \end{array}\right\} \quad \text{for} \quad \left\{\begin{array}{c} \rho^k = \pi \\ \rho^k > \pi \end{array}\right\} \tag{47}$$

Such a contract ensures that a weak bank will not experience a run. Moreover, among all run-proof contracts, the contract in (46) - (47) cannot be improved upon by another run proof contract (Proposition 6). The contract in (46) - (47) will be optimal if the following condition is satisfied (Proposition 6)

$$u\left(\min\{c_{1S}^*,\, \hat{c}_2\}\right) - u(\hat{c}_2) \leq (1-\pi)\left(u(\hat{c}_2) - u(\hat{c}_1)\right) \tag{48}$$

This condition ensures that bank-k cannot deliver strictly higher welfare to its investors by deviating to a contract that leads to a run when the bank has a weak fundamental. The payments in resolution $(\hat{c}_1, \hat{c}_2)$ are determined from:

$$\theta \min\{c_{1S}^*,\, \hat{c}_2\} + (1-\theta)\left[\pi\hat{c}_1 + (1-\pi)\frac{\hat{c}_2}{R}\right] = 1 - \tau - \bar{\sigma} + \hat{b}$$

$$u'(\hat{c}_1) = Ru'(\hat{c}_2) = v'\left(\tau^* - n\hat{b}\right)$$

The existence of the equilibrium without a run is numerically established in Figure 2 . The same figure also establishes multiple equilibria for some parameter values. We collect the allocation in the run equilibrium in the following vector:

$$\mathrm{D}^{\mathrm{NR}} \equiv \left(\tau^*,\, c_{1S}^*,\, c_{2S}^*,\, \hat{c}_1^{NR},\, \hat{c}_2^{NR},\, \hat{b}^{NR}\right)$$

□

Proposition 11.

Proof. Denote with $CE(n)$ the constrained efficient allocation for given n:

$$CE(n) \equiv (c_{1S}^*(n),\, c_{2S}^*(n),\, c_{1W}^*(n),\, c_{2W}^*(n),\, b^*(n),\, \tau^*(n))$$

Similarly, denote with $R^{DE}(n)$ the allocation in the decentralized economy where weak banks experience a run from their investors:

$$R^{DE}(n) \equiv \left(c_{1S}^*(n),\, c_{2S}^*(n),\, \hat{c}_1(n),\, \hat{c}_1(n),\, \hat{b}(n),\, \tau(n),\, \bar{c}(n)\right)$$

where the tax rate $\tau^{DE}(n) = \tau^{CE}(n)$ is fixed to its level in the constrained efficient allocation and the cap in early payments is fixed to $\bar{c}(n) = c_{1S}^*(n)$. As $n \to 0$ we have

$$u'(\hat{c}_1(n)) = v'(\tau^*(n) - nb^*(n)) \to v'(\tau(n))$$

$$u'(c_{1S}^*(n)) \to v'(\tau(n))$$

Therefore $c_{1S}^*(n) \to \hat{c}_1(n)$ and $c_{2S}^*(n) \to \hat{c}_2(n)$. Then, since $c_{1S}^*(n) < c_{2S}^*(n)$ we obtain that there exist $\bar{n} > 0$ such that for each $n < \bar{n}$ we have $c_{1S}^*(n) < \hat{c}_2(n)$. In other words, the run equilibrium does not exist for values of n that are sufficiently close to zero (i.e. those below $\bar{n}$). □

Proposition 12.

Proof. The objective of the government in period 0 is to choose $\bar{c}$ - the cap on period 1 payments in the bad aggregate state - in order to maximize the sum of investors' utilities from the perspective of that period.

$$(1-q)\left[V_S(1-\tau)+v(\tau)\right]+q\left\{\begin{array}{c}(1-n)\left[\pi u(\bar{c})+(1-\pi)V_S\left(1-\tau-\pi\bar{c}\right)\right]\\ +n\left[\theta u(\bar{c})+(1-\theta)V_S\left(1-\tau-\bar{\sigma}-\theta\bar{c}+b(\bar{c})\right)\right]\\ +v\left(\tau-nb(\bar{c})\right)\end{array}\right\} \quad (49)$$

where $V_S(.)$ is the the per capita sum of utilities for the remaining investors in a sound bank as a function of the available resources, $V_W(.)$ is the per capita sum of utilities for the remaining investors in weak a bank as a function of the available resources (which, in the case of a weak bank, might involve a bailout). The program in (49) is equivalent to choosing $\bar{c}$ in order to maximize the term in the curly bracket in the above expression. The first order condition characterizing the optimal choice of $\bar{c}$ is given in (**??**) and can be re-written as:

$$u'\left(\bar{c}\right)-Ru'\left(\frac{R(1-\tau-\pi\bar{c})}{1-\pi}\right)=\frac{n\theta}{(1-n)\pi}\left[v'\left(\tau-nb(\bar{c})\right)-u'\left(\bar{c}\right)\right] \quad (50)$$

Evaluating (50) at the constrained efficient choice of period 1 payments in sound banks c_{1S}^*:

$$0=u'\left(c_{1S}^*\right)-Ru'\left(\frac{R(1-\tau-\pi c_{1S}^*)}{1-\pi}\right)<\frac{n\theta}{(1-n)\pi}\left[v'\left(\tau-nb(c_{1S}^*)\right)-u'\left(c_{1S}^*\right)\right]$$

Similarly, evaluating (50) at the constrained efficient choice of period 1 payments in weak banks c_{1W}^*:

$$u'\left(c_{1W}^*\right)-Ru'\left(\frac{R(1-\tau-\pi c_{1W}^*)}{1-\pi}\right)>\frac{n\theta}{(1-n)\pi}\left[v'\left(\tau-nb(c_{1W}^*)\right)-u'\left(c_{1W}^*\right)\right]=0$$

Therefore, the optimal choice of the cap in early payments $\bar{c}^*$ must satisfy $c_{1W}^*<\bar{c}^*<c_{1S}^*$. This establishes the first part of the Proposition 10, namely $\bar{c}^*<c_{1S}^*$.

Next, by examining the first order condition in (50) we see that the left hand side is independent of n, $\bar{\sigma}$ and θ and decreasing in $\bar{c}$. Next, defining the function f as follows

$$f(\bar{c},n,\bar{\sigma},\theta)\equiv\frac{n\theta}{(1-n)\pi}\left[v'\left(\tau-nb(c_{1W}^*)\right)-u'\left(c_{1W}^*\right)\right]$$

Holding $(n,\bar{\sigma},\theta)$ fixed we have $\dfrac{\partial f(\bar{c},n,\bar{\sigma},\theta)}{\partial\bar{c}}<0$. On the other hand:

$$\frac{\partial f(\bar{c},n,\bar{\sigma},\theta)}{\partial n}>0,\quad \frac{\partial f(\bar{c},n,\bar{\sigma},\theta)}{\partial\bar{\sigma}}>0\quad\text{and}\quad\frac{\partial f(\bar{c},n,\bar{\sigma},\theta)}{\partial\theta}>0$$

That is, the right hand side of (50) is increasing in n, $\bar{\sigma}$ and θ and decreasing in $\bar{c}$. Therefore, we obtain that $\bar{c}^*$ is decreasing in n, $\bar{\sigma}$ and θ. Moreover, $\bar{c}^*\rightarrow c_{1S}^*$ as either $n\rightarrow 0$ or

$\theta \to 0$. □

Proposition 13.

Proof. In period 0, the government chooses the tax rate τ and the cap $\bar{c}$ in order to maximize the program in (49). The optimal choice of the tax rate τ^D must satisfy the first order condition in (**??**), which can be equivalently expressed as:

$$v'\left(\tau^D\right) = u'\left(c_{1S}\left(\tau^D\right)\right) - \frac{q(1-n)}{1-q}\left[v'\left(\tau^D - nb^D\left(\tau^D, \bar{c}^*\right)\right) - Ru'\left(\frac{R(1-\tau^D-\pi\bar{c}^*)}{1-\pi}\right)\right] \tag{51}$$

where $c_{1S}(\tau^D)$ the period 1 payment in the sound banks (in the good aggregate state), $b^D\left(\tau^D, \bar{c}^*\right)$ is the per capita bailout payment to each weak bank in the bad aggregate state and $\bar{c}^*$ is the cap on early payments in the bad aggregate state - characterized by (50).

On the other hand, the optimal tax rate in in the constraint efficient allocation is determined from the following first order condition:

$$v'\left(\tau^*\right) = u'\left(c_{1S}^*\right) - \frac{q(1-n)}{1-q}\left[v'\left(\tau^* - nb^*\right) - Ru'\left(c_{2S}^*\right)\right]$$

where b^* is the per capita bailout payment to each weak bank in the constrained efficient allocation. From Proposition 10 we have $c_{1W}^* < \bar{c}^* < c_{1S}^*$ which yields:

$$b^* < b^D\left(\tau^D, \bar{c}^*\right) \quad \Leftrightarrow \quad \tau^* - nb^D\left(\bar{c}^*, \tau^*\right) < \tau^* - nb^*$$

Combining this inequalities, we obtain that in the first order condition in (51) will not hold unless the government sets a higher tax rate in the decentralized economy, that is, $\tau^D > \tau^*$. □