Bailouts, Bail-ins and Banking Crises*

Todd Keister
Rutgers University
todd.keister@rutgers.edu

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

October 21, 2017

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.

*Preliminary and incomplete draft. 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.

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 contractual environment have been completely successful. We ask whether

[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).

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 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

[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.

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's objective is to maximize the sum of 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. After collecting taxes, the government in our model moves two more times: first as a *banking regulator* and then as a *fiscal authority*. We describe the actions of the fiscal authority first.

Fiscal authority: 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 - acting as a fiscal authority - decides how to allocate its tax revenue between bailout payments to banks and the public good. Banks that receive a bailout from the government are immediately placed in *resolution* and all subsequent payments made by these banks are chosen by the government. The parameter θ thus measures how quickly the government can collect bank-specific information during a crisis and respond to this information. Once the public good has been provided, the government no longer has access to any resources and there will be no further bailouts.

Banking regulator: the banking regulator can restrict the set of payments each bank is allowed to make to its investors during the first θ withdrawals. These restrictions on payments are imposed before the start of period 1 withdrawals and remain in force until the subsequent intervention of the fiscal authority whose behavior was described above.

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. The government regulator announces payment restrictions that will be in place in period 1 as a function of the aggregate state and then banks choose their contracts. At the beginning of period 1, each investor observes her own preference type

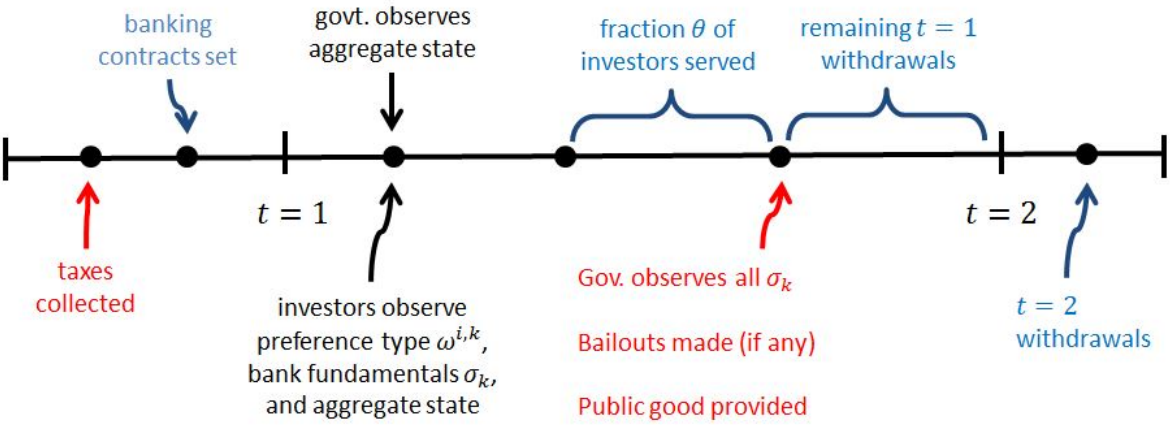

Figure 1: Timeline

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. Payments to withdrawing investors must be consistent with the existing regulatory restrictions. 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 to a crisis would be captured in our model by an increase in the parameter θ.

3 The 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

[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.

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

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

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,

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

[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.

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

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

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

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

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]

Proposition 1. *The 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 efficient allocation satisfies*

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

[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.

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 within a bank

In this section, we begin our analysis of the equilibrium of the game defined in Section 2. The analysis proceeds as follows: in Section 4 we fix the policy of the banking regulator and characterize the outcome within a bank, taking as given the strategies of all other banks and their investors. In Section 5 we study the joint determinant of equilibrium across all banks - again holding fixed the policy of the banking regulator. Finally, in Section 6 we derive the optimal regulatory policy.

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.[9]

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 the aggregate state, the bank's

[9] The analysis of the optimal choice of the tax rate will not contribute significantly to the analysis and is omitted.

fundamental state $\sigma_k \in \{0, \bar{\sigma}\}$ and the fraction $\rho^k \in [\pi, 1]$ of its investors who choose to withdraw early. In this section we focus on the form of banking contracts conditional on the bad aggregate state where fraction n of the banks have weak fundamentals. It will become clear that the optimal banking contracts in the good aggregate state can be obtained as a special case of the analysis in this section.

Since 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.[10] Let c_1^k denote the payment made by the bank to each investor who withdraws in period 1. In period 2, the bank 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 which is a mapping from the bank-specific state ρ^k and σ_k to the menu of payment options provided by the regulator X_k. That is,

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

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.[11]

Banking Regulation. The banking regulator presents each bank with a menu of payment options to choose from. The menu of payment options available to bank-k in the bad aggregate state will be denoted X_k. For example, if bank-k is presented with $X_k \equiv \{0, \bar{c}\}$, then the bank can only choose between a strict suspension of payments $c_1^k = 0$ or setting $c_1^k = \bar{c}$. On the other hand, if $X_k = [0, \bar{c}]$ then bank-k allowed to set any payment up to an upper bound equal to $\bar{c}$. Notice that the menu of payment options offered to bank-k cannot be made contingent on the bank-specific state, σ_k and ρ^k, since this information is not available to the banking regulator.

Also, notice that any regulatory policy implementing the efficient outcome from Section 3 must necessarily permit a weak bank the option to set c_{1W}^* and a sound bank the option to set c_{1S}^*. In this section and in section 5 we assume that the menu of payments options offered to each bank in the bad aggregate state satisfies the following *Baseline Case*:

[10] 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.

[11] 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.

$$\{0,\, c_{1W}^{*},\, c_{1S}^{*}\} \subseteq X_k \tag{10}$$

where c_{1W}^{*} and c_{1S}^{*} were derived in Section 3 and the option to set 0 allows each bank to impose a strict suspension of payments. In Section 6 we analyze the optimal regulatory policy and show that under some circumstances the banking regulator might decide to **remove** either c_{1W}^{*} or c_{1S}^{*} (or both) from the menu of options available to banks. The rationale for doing so will become clear once we characterize the outcome where both of these options are made available to each bank.

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. For given σ_k, we say investor i in bank k follows (i) the *no-run strategy* if he withdraws in period 1 only if impatient, that is, $y_k^i\left(\omega_k^i,\, \sigma_k\right) = \omega_k^i$ for $\omega_k^i \in \{0, 1\}$ and (ii) the *run strategy* if he withdraws in period 1 even if patient, that is, $y_k^i\left(\omega_k^i,\, \sigma_k\right) = 0$ 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 strategy in that profile, which we denote

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

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

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

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.[12] 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 char-

[12]This result follows, in part, from the assumption that there are a continuum of locations and, hence, the actions taken at one bank have no effect on aggregate variables or on the behavior of the government toward other banks.

acterize 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.[13] 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 to withdraw at $t = 2$ receive an even share of the bank's assets, $c_2^k(0,\, \rho^k)$, which is implicitly defined by

$$p^k c_1(0,\, \rho^k) + (1-\rho^k)\frac{c_2(0,\, \rho^k)}{R} = 1-\tau. \tag{11}$$

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

$$\rho^k = \left\{ \begin{array}{c} \pi \\ \in [\pi, 1] \\ 1 \end{array} \right\} \quad \text{as} \quad c_1(0,\, \rho^k) \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} c_2(0,\, \rho^k). \tag{12}$$

We refer to (12) as the *implementability constraint*. If a triple (ρ_S, c_{1S}, c_{2S}) satisfy both (11) and (12), 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.

Proposition 3. *If the allocation (ρ_S, c_{1S}, c_{2S}) satisfies both (11) and (12), 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

$$V_S\,(\rho_S,\, c_{1S}) \equiv \rho_S u(c_{1S}) + (1-\rho_S)u(c_{2S}) \tag{13}$$

subject to the feasibility constraint in (11) the implementability constraint in (12) and the regulatory restriction on banking contracts in (10) which can simply be expressed as requiring $c_{1S} \in X_k$. The next result shows that whenever c_{1S}^* is a payment bank-k

[13] 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.

is allowed to set the equilibrium allocation within a sound bank is the same as in the efficient allocation.

Proposition 4. *When bank k has sound fundamentals and $c_{1S}^* \in X_k$ 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 $c_{1S} = c_{1S}^*$.*

In other words, resources are always allocated efficiently within a sound bank and investors never run on these banks.[14] There are many banking contracts that implement the desired allocation, one of which is

$$c_1^k\left(0,\, \rho^k\right) = \left\{ \begin{array}{c} c_{1S}^* \\ 0 \end{array} \right\} \quad \text{if} \quad \left\{ \begin{array}{c} \rho^k = \pi \\ \rho^k > \pi \end{array} \right\}. \tag{14}$$

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.[15] 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. 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.

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,

$$W \equiv \{k \in [0,1] \text{ s.t } \sigma_k = \bar{\sigma}\}\,. \tag{15}$$

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.

[14] The allocation within sound banks in case c_{1S}^* might no longer be an option provided by the regulator will be characterized in Section 6.

[15] 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 (14) makes this point in a particularly clean way.

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

$$\hat{\psi}_k = \frac{(1-\tau)(1-\bar{\sigma}) - \theta c_1^k(\bar{\sigma},\, \rho_k) + b_k}{1-\theta}, \tag{16}$$

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:

$$\hat{V}\left(\hat{\psi}_k;\, \hat{\rho}_k\right) \equiv \max_{\hat{c}_1^k,\, \hat{c}_2^k} \; (1-\theta)\left(\hat{\rho}_k u\left(\hat{c}_1^k\right) + (1-\hat{\rho}_k) u\left(\hat{c}_2^k\right)\right) \tag{17}$$

subject to the feasibility constraint

$$\hat{\rho}_k \hat{c}_1^k + (1-\hat{\rho}_k)\frac{\hat{c}_2^k}{R} \leq \hat{\psi}_k \tag{18}$$

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

$$u'\left(\hat{c}_1^k\right) = Ru'\left(\hat{c}_2^k\right) = \hat{\mu}(\hat{\psi}_k;\, \hat{\rho}_k), \tag{19}$$

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

$$\max_{\{b^k\}_{k\in W}} \int_W \hat{V}\left(\hat{\psi}_k;\, \hat{\rho}_k\right) dk + v\left(\tau - \int_W \hat{b}_k dk\right) \tag{20}$$

The first-order condition for this problem is

$$\hat{\mu}(\hat{\psi}_k;\, \hat{\rho}_k) = v'\left(\tau - \int_W \hat{b}_k dk\right) \text{ for all } k. \tag{21}$$

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

$$\hat{\rho}_k \hat{c}_1 + (1-\hat{\rho}_k)\frac{\hat{c}_2}{R} = \frac{(1-\tau)(1-\bar{\sigma}) - \theta c_1^k(\bar{\sigma},\, \rho_k) + b_k}{1-\theta}. \tag{22}$$

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, $c_1^k(\bar{\sigma},\, \rho_k)$. Together with the first-order condition (19), 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 within a weak bank. A fraction θ of a weak bank's investors will receive the amount specified by the contract, $c_1^k(\bar{\sigma},\, \rho_k)$, 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

$$\rho^k = \left\{ \begin{array}{c} \pi \\ \in [\pi, 1] \\ 1 \end{array} \right\} \quad \text{if} \quad c_1^k(\bar{\sigma},\, \rho_k) \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} \hat{c}_2. \tag{23}$$

In choosing a contract, the bank recognizes that its investors will behave in accordance with (23), 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 (23), 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

$$V_W(\rho_W, c_{1W}) \equiv \theta u(c_{1W}) + (1-\theta)\left[\hat{\rho}_W u(\hat{c}_1) + (1-\hat{\rho}_W) u(\hat{c}_2)\right] \tag{24}$$

subject to the implementability constraint for weak banks (23), the regulatory restriction on contracts in (10) namely $c_{1W} \in X_k$, and the relationship

$$\hat{\rho}_W \equiv \frac{\pi}{1-\theta}\left(\frac{\rho_W - \theta}{\rho_W}\right). \tag{25}$$

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 (24) is clearly increasing in 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 (23) 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 $\hat{\rho}_W$ is an increasing function of ρ^k and the second term in the objective function (24) is strictly decreasing in $\hat{\rho}_W$.

Equilibrium within a weak bank. Before characterizing the equilibrium outcome within a weak bank, we introduce some useful labels. Denote with $c_1^k(\mathrm{H})$ the highest payment among those permitted by the regulator for bank-k. That is, $c_1^k(\mathrm{H}) \in X_k$ such that

$$c_1^k(\mathrm{H}) \geq c_1 \quad \text{for any } c_1 \in X_k. \tag{26}$$

Also denote with $c_1^k(\mathrm{H}\,|\mathrm{NR})$ the highest payment among those that are *both* regulatory permitted and avoid a run within the bank i.e. those that are less or equal to $\hat{c}_2$. That is, $c_1^k(\mathrm{H}\,|\mathrm{NR}) \in X_k$ and $c_1^k(\mathrm{H}\,|\mathrm{NR}) \leq \hat{c}_2$ such that

$$c_1^k(\mathrm{H}\,|\mathrm{NR}) \geq c_1 \quad \text{for any } c_1 \in X_k \text{ such that } c_1 \leq \hat{c}_2. \tag{27}$$

Notice that $c_1^k(\mathrm{H}\,|\mathrm{NR})$ corresponds to an optimal *run-proof* allocation from the perspective of bank-k. Given that the regulatory policy in (10) allows each bank to set $c_1^k = 0$, there exist at least one run-proof contract available to bank-k and therefore $c_1^k(\mathrm{H}\,|\mathrm{NR})$ will be well defined.

Proposition 6. *The solution to the program of maximizing (24) subject to (23), (25) and $c_{1W} \in X_k$ has the following property: the early payment chosen by the bank will either be set as high as the regulator allows $c_{1W} = c_1^k(\mathrm{H})$, or will be set to the highest value that is both permitted and prevents a run within the bank, $c_{1W} = c_1^k(\mathrm{H}\,|\mathrm{NR})$.*

Note that (24) implies that setting $c_{1W} = c_1^k(\mathrm{H})$ will be optimal whenever a weak bank is able to *both* pay the maximum amount allowed to the first θ investors without generating a run among its investors, that is, $c_1^k(\mathrm{H}) \leq \hat{c}_2$. On the other hand, if $c_1^k(\mathrm{H}) > \hat{c}_2$ then the bank would still choose to set its payment to the highest value permitted by the regulator - run notwithstanding - whenever the following condition is satisfied:

$$V_W\left(c_1^k(\mathrm{H}),\, 1\right) > V_W\left(c_1^k(\mathrm{H}\,|\mathrm{NR}),\, \pi\right).$$

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.

Next, by observing that the inequality above is equivalent to

$$u\left(c_1^k(\mathrm{H})\right) - u\left(c_1^k(\mathrm{H}\,|\mathrm{NR})\right) > (1-\pi)\left(u\left(\hat{c}_2\right) - u\left(\hat{c}_1\right)\right)$$

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

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

(i) If $u\left(c_1^k(\mathrm{H})\right) - u\left(c_1^k(\mathrm{H}\,|\mathrm{NR})\right) < (1-\pi)\left(u\left(\hat{c}_2\right) - u\left(\hat{c}_1\right)\right)$, 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 $c_{1W} = c_1^k(\mathrm{H}\,|\mathrm{NR})$.

(ii) If $u\left(c_1^k(\mathrm{H})\right) - u\left(c_1^k(\mathrm{H}\,|\mathrm{NR})\right) > (1-\pi)\left(u\left(\hat{c}_2\right) - u\left(\hat{c}_1\right)\right)$, 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 $c_{1W} = c_1^k(\mathrm{H})$.

(iii) If $u\left(c_1^k(\mathrm{H})\right) - u\left(c_1^k(\mathrm{H}\,|\mathrm{NR})\right) = (1-\pi)\left(u\left(\hat{c}_2\right) - u\left(\hat{c}_1\right)\right)$, the withdrawal game in bank k has multiple equilibria, one with $\rho_W = \pi$ and $c_{1W} = c_1^k(\mathrm{H}\,|\mathrm{NR})$ and another with $\rho_W = 1$ and $c_{1W} = c_1^k(\mathrm{H})$.

This result shows that, outside of a knife-edge case in *(iii)*, 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.

5 Equilibrium across banks

In this section we being our investigation of 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. Moreover, the regulatory policy of the government is fixed to the Baseline Case in (10), namely both c^*_{1W} and c^*_{1S} is made available to each bank. The next result shows that efficiency is no longer a property of the equilibrium allocations.

Proposition 8. *The equilibrium is inefficient whenever the set of weak banks is identified with a delay.*

A necessary condition for efficiency is that all weak banks self-select into c^*_{1W} . So, suppose all other weak banks set c^*_{1W} and consider a given weak bank. In order for such an allocation to be sustained in equilibrium each individual weak bank must have no incentive to unilaterally deviate by choosing a different payment from the menu of options available. By setting c^*_{1W}, the sum of utilities for the investors within the bank will be given by:

$$\theta u\left(c^*_{1W}\right) + (1-\theta)\left[\hat{\rho}_W u\left(c^*_{1W}\right) + (1-\hat{\rho}_W)\, u\left(c^*_{2W}\right)\right] \tag{28}$$

where $\hat{\rho}_W$ will be equal to $(\pi-\theta)/(1-\theta)$ as obtained from the relation in (25). On the other hand, by setting c^*_{1S} (recall that both c^*_{1W} and c^*_{1S} belong to (10) and therefore can be chosen by each bank) the sum of utilities for the investors within the bank becomes:

$$\theta u\left(c^*_{1S}\right) + (1-\theta)\left[\hat{\rho}^d_W u\left(c^*_{1W}\right) + \left(1-\hat{\rho}^d_W\right) u\left(c^*_{2W}\right)\right] \tag{29}$$

where $\hat{\rho}^d_W$ is obtained from (25) and denotes the fraction of remaining impatient investors after the first θ withdrawals in this case. Notice that for any value of $\hat{\rho}^d_W$ the expression in (29) will be strictly greater than the expression in (28). Therefore setting c^*_{1W} is **not** incentive compatible for weak banks. In other words, a weak bank can deliver a strictly higher utility to its investors by pretending to be sound and setting c^*_{1S} and as a result will choose not to self-select into c^*_{1W}.

Summarizing: the allocation within weak banks will not be efficient, whenever c^*_{1S} is an option. At the same time, the allocation within sound banks cannot be efficient unless c^*_{1S} is an option. Jointly, these two claims imply that the economy-wide allocation in the bad aggregate state cannot be efficient, regardless of the regulatory policy followed by the government.

5.1 Fragility

In this section, we show that, in addition to being inefficient, for some parameters the equilibrium will be consistent with a welfare decreasing run by investors on the weak banks. Specifically, in the *run equilibrium*, all investors in weak banks attempt to withdraw at $t = 1$. A fraction θ of these investors successfully withdraw before the government observes $\sigma_k = \bar{\sigma}$ and places the bank into resolution. The equilibrium of the full model will, in some cases, involve a run by investors on weak banks as shown in the next result.

Proposition 9. *Suppose the regulatory policy is such that both c^*_{1W} and c^*_{1S} are among the choices available to each bank in the bad aggregate state. Then 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.*

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[16]

$$u(c_1^{i,k} + \omega^{i,k} c_2^{i,k}) = \frac{\left(c_1^{i,k} + \omega^{i,k} c_2^{i,k}\right)^{1-\gamma} - 1}{1-\gamma} \qquad \text{and} \qquad v(g) = \delta \frac{g^{1-\gamma}}{1-\gamma}. \tag{30}$$

The regulatory policy is the following: the banking regulator imposes that the payment made by each bank does not exceed an upper limit of $\bar{c}$. A natural choice of this upper limit is to set it at least equal to the payment made by sound banks in the efficient allocation $\bar{c} \geq c^*_{1S}$ since setting a limit below c^*_{1S} prevents sound banks from allocating resources efficiently. Henceforth we set $\bar{c} = c^*_{1S}$ and thus, the menu of payment options available to each bank in both aggregate states will be given by $X = [0, c^*_{1S}]$.

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 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

[16] 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 efficient value from section 3.

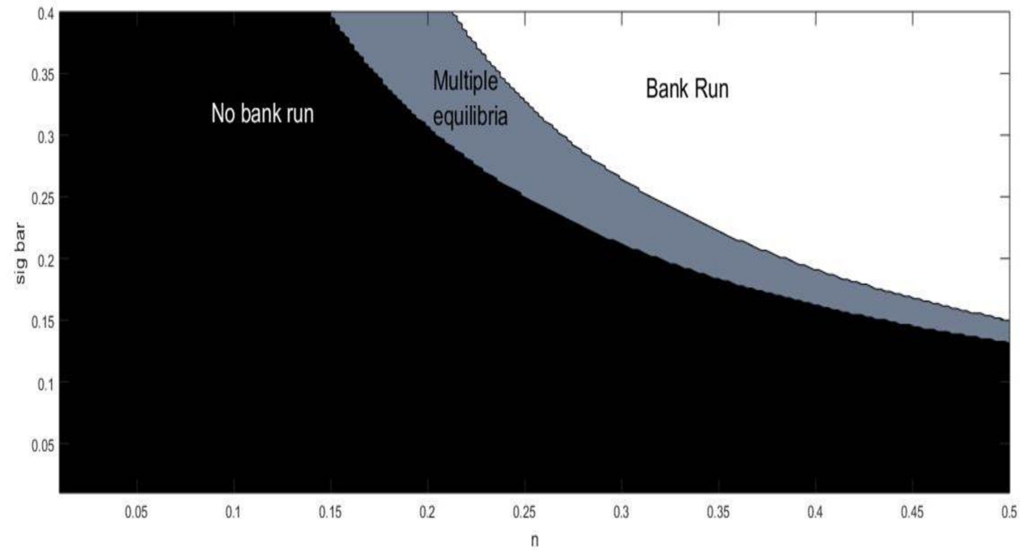

Figure 2: Equilibrium with a bank run

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.[17]

Runs must be systemic. The pattern in Figure 2 suggests that a run can only occur when the number of weak banks is sufficiently large. The following proposition formalizes this result.

Proposition 10. *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. Thus, 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 or a small group of banks.

[17]The fact that each bank is allowed to choose among many (a continuum) of options in the example above is not critical. Alternately we could assume that, instead of imposing an upper limit, each bank is allowed to choose only between three payment options 0, c^*_{1W} and c^*_{1S}. The results in this case will be qualitatively similar to those in Figure 2.

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.

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 United States, 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.

6 Optimal regulation

In the previous sections, the period 1 payments permitted to banks by the regulator were assumed to always include both c^*_{1S} and c^*_{1W}. Supposing that the regulator does not explicitly prevent a bank from setting either of these payments seemed like a reasonable approach to regulation, especially given the fact that any regulatory policy ensuring efficiency must necessarily allow each bank to self-select into c^*_{1S} or c^*_{1W}. As we established in Proposition 8, however, the allocation where each weak bank self-selects into c^*_{1W} is not incentive compatible and therefore cannot be sustained in equilibrium. Furthermore, for some parameters the economy with regulation as in the baseline case in (10) can have as a unique equilibrium a run on all weak banks (see Proposition 9 or Figure 2).

In this section we begin our analysis of the optimal regulatory policy. Specifically, we allow the banking regulator to freely choose the menu of payment options to be made available to each bank as a function of the aggregate state. We ask the question: how should the banking regulator select the menu of options offered to each bank in the good and in the bad aggregate state respectively? First, suppose that the aggregate state is good in which case the regulator knows that all banks must be sound. Then an optimal regulatory policy is to offer each bank a choice between two options only: 0 and c^*_{1S}. As before, the option of picking 0 is there to ensure that banks can write run proof contracts with their investors, whereas the option c^*_{1W} is not on the table since there are no weak banks. Observe that our decision to focus on the bad aggregate state in the exposition of Sections 4 and 5 was without loss of generality. Indeed, if the aggregate state turns out to be good then the regulator infers all banks must be sound in which case an appropriately choosing regulatory policy as explained above can ensure that the resulting allocation is efficient and runs never occur. Next we turn to the optimal regulation when fraction of the banks are weak.

6.1 Optimal regulation when some banks are weak

Our goal in this section is to derive optimal regulation in the bad aggregate state. First, suppose the regulator were able differentiate weak from sound banks from the very beginning of period 1 and consider a regulatory policy where each sound bank is offered a choice between 0 and c^*_{1S} and each weak bank is offered a choice between 0 and c^*_{1W}. That is, banks are allowed to suspend payments on their own. In this case a run can never occur and the equilibrium outcome corresponds to the efficient one from Section 3. The reason is that the best strategy available to a weak bank is to set $c^k_1 = c^*_{1W}$ whenever $\rho^k = \pi$ and to completely freeze payments otherwise (and similarly for a sound bank, but with c^*_{1S} in place of c^*_{1W}). Thus the regulatory ability to separate the weak from the sound banks allows the government to design regulation which restores economic efficiency and ensures that banks would enact contracts which are run proof.

For the rest of this section we focus on the interesting case where the set of weak banks is not known initially to the regulator and the aggregate state is bad. For a given menu of payment options in the bad aggregate state X the equilibrium will be characterized by the allocation within sound banks ρ_S and c_{1S}, the allocation within weak banks ,ρ_W and c_{1W}, and the bailout and resolution policy of the government $\hat{c}_1$, $\hat{c}_2$ and $\hat{b}$. We collect all these variables in a vector E which we refer to as the *economy-wide* allocation:

$$E \equiv \left(\rho_W, \rho_S, c_{1S}, c_{2S}, c_{1W}, \hat{c}_1, \hat{c}_2, \hat{b}\right) \tag{31}$$

We make the simplifying observation that focusing on menu of payment options which are *3-tuples* such as $X = \{0,\, c_{1W},\, c_{1S}\}$ (*or 2-tuple* whenever $c_{1W} = c_{1S}$) is in fact without loss of generality.[18]

Program of the regulator when weak banks are unknown. The program of the regulator in the bad aggregate state is to choose c_{1W} and c_{1S} in order to maximize the sum of all investors utilities taking into account the incentives faced by the banks and their investors and the bailout intervention of the fiscal authority after the first θ withdrawals.

Alternatively, we can formulate the problem of the banking regulator as suggesting an

[18]The reason is that augmenting the payment menu X with an additional element creates additional possibilities for unilateral deviations by individual banks and therefore the regulatory policy would have to satisfy a greater number of incentive compatibility constraints. Indeed, consider a given set $X = \{0,\, c_{1W},\, c_{1S}\}$ such that, in equilibrium, all weak banks choose c_{1W} and all sound banks choose c_{1S}. Now, suppose a new element z is added to those already existing in X. If the switch to z is undesirable from regulatory perspective then adding z can only reduce equilibrium equilibrium welfare. On the other hand, if a switch to z is desirable either for the weak or for the sound banks (or for both) then a better regulatory policy will be replacing z with the element(s) in X it is superior to (from regulatory perspective). The reason is that keeping both the new and the old elements in X creates incentives for deviations which - while optimal for individual banks - might be undesirable from regulatory perspective.

allocation to each bank. Observe that the regulator cannot control investors' withdrawal choices and therefore the allocation within each bank must satisfy the constraints imposed by investors' withdrawal strategies. That is the allocation must satisfy (11) - (12) if the bank is sound and (23) if the bank is weak. In addition, the regulator cannot control the bailout policy of the fiscal authority and therefore after the first θ withdrawals the continuation allocation within each weak bank, $\hat{c}_1$ and $\hat{c}_2$, must satisfy the constraints imposed by the ex-post optimal bailout intervention in (19) - (22). Finally, the regulator does not observe bank-specific states, ρ_k and σ_k, and hence must ensure that weak and sound banks self-select into their corresponding payment. Specifically, if a weak bank reports to be sound, it will be allowed to set c_{1S} and therefore any allocation, $\tilde{\rho}_W$ and c_{1S}, satisfying the implementability constraint for weak banks in (23) can be attained as an equilibrium outcome within this bank. The incentive compatibility constraint for a weak bank can therefore be expressed as follows:

$$V_W(\rho_W, c_{1W}) \geq V_W(\tilde{\rho}_W, c_{1S}) \qquad \text{for all } \tilde{\rho}_W \in [\pi, 1] \text{ for which (23) holds} \tag{32}$$

If the above were not true, i.e. if there exist $\tilde{\rho}^*_W$ such that $V_W(\rho_W, c_{1W}) < V_W(\tilde{\rho}^*_W, c_{1S})$, then a weak bank will be able to increase the sum of utilities for its investors by setting c^*_{1S} and then implementing the allocation corresponding to $\tilde{\rho}^*_W$ and c_{1S}. Finally, the incentive compatibility constraint for a sound banks is provided below:

$$V_S(\rho_S, c_{1S}) \geq V_S(\tilde{\rho}_S, c_{1W}) \qquad \text{for all } \tilde{\rho}_S \in [\pi, 1] \text{ for which (12) holds} \tag{33}$$

The next result summarizes the conditions under which the economy-wide allocation can be obtained as an equilibrium of a given regulatory policy.

Proposition 11. *Assume that in the bad aggregate state the banking regulator offers each bank a choice between* 0, c_{1W} *and* c_{1S}. *Such a regulatory policy is without any loss of generality. Furthermore, the regulatory policy* $X = \{0, c_{1W}, c_{1S}\}$ *can result in the economy-wide allocation* E *as an equilibrium if and only if* E *jointly satisfies (11) - (12), (32) - (33) and (19) - (22)*

In light of the this proposition we can recast the problem of choosing the optimal regulatory policy in the bad aggregate state as one of directly choosing the economy-wide allocation E in (31) to maximize the sum of all investors utilities:

$$(1-n)V_S(\rho_S, c_{1S}) + nV_W(\rho_W, c_{1W}) + v\left(\tau^* - n\hat{b}\right) \tag{34}$$

subject to (11) - (12), (32) - (32) - (33) and (19) - (22). Next, we introduce some notation that will be useful when characterizing optimal regulation in the bad aggregate state.

Systemic regulatory bail-in. A *systemic bail-in policy* is one where the regulator gives each bank just two options: either suspend payments completely $c_1^k = 0$ or set $c_1^k = \bar{c}$ such that $\bar{c} < c_{1S}^*$. The benefit of such a policy is that the incentive for a weak banks to pretend to be sound is addressed by requiring all banks to enact the same bail-in in the bad aggregate state. Such a policy comes with a trade-off, however, since it imposes a cost on the sound banks which will be forced into an unnecessary bail-in. Proposition 12 below shows that it can nevertheless be optimal for the regulator to require all banks to bail-in their investors to $\bar{c}$ or suspends all payments completely.

Note that in this environment a systemic 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.

Selective regulatory bail-in. A *selective bail-in policy* is one where the regulator offers each bank three options: either suspend payments completely $c_1^k = 0$ or choose between two different payments c_{1W} and c_{1S} such that $c_{1W} < c_{1S}$. 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}. Second, 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. It is worth pointing out that 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.

The next result shows that the optimal bail-in policy can be selective or systemic depending in parameter values.

Proposition 12. *Weak banks are never required to fully bail-in their investors. For some parameters, the optimal policy is a systemic bail-in. For other parameters, the optimal policy is a selective bail-in.*

Proposition 12 is established in Figure 3 showing the optimal bail-in policy contingent 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 $R = 1.5$, $\pi = \theta = 0.5$, $\delta = 0.5$, $\gamma = 6$, $q = 0.05$. For all combinations of n and $\bar{\sigma}$ in the black region of Figure 3, the optimal bail-in policy is to trigger a systemic 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 partial bail-in.

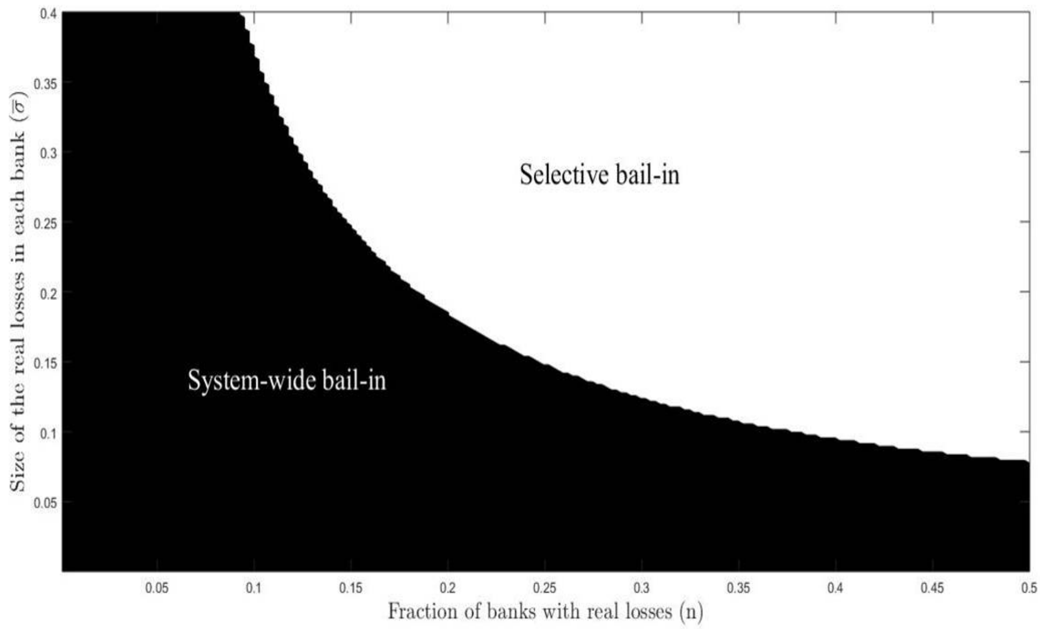

Figure 3: Optimal regulation

6.2 Remarks

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.

In addition, note 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 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.[19]

Finally, suppose the government can commit to *strict no-bailout rule.* In this case, runs can no longer occur as part of equilibrium. The reason is that in our environment each bank 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. Furthermore, the question of when weak banks are identified by the government becomes irrelevant in the absence of bailouts since each all weak banks will face the correct incentives when it comes to bailing in their investors. At the same time, a no-bailout policy would dispense with the socially valuable ex-post insurance function of the government's bailouts (recall that weak banks experience real losses in our model). 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.

7 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)

[19] 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.

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: Proof of selected Propositions

Proposition 1.

Proof. We need to show that sound banks do not receive bailouts and the consumption of investors in sound banks is independent of the aggregate stat. Observe that the following condition can be shown to hold:

$$b_S^* = 0 \quad \Leftrightarrow \quad (c_{10}^*, c_{20}^*) = (c_{1S}^*, c_{2S}^*) \tag{35}$$

This follows from the feasibility constraints in (2) and (3) and the first order conditions in (6). Therefore it will be sufficient to show $b_S^* = 0$. 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 (35) we obtain that $(c_{10}^*, c_{20}^*) = (c_{1S}^*, c_{2S}^*)$, namely the payment profile in sound banks is independent of the aggregate state. $\square$

Proposition 2.

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

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

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

$$c_{1S}^* \leq c_{1W}^* \quad \text{and} \quad c_{2S}^* \leq c_{2W}^* \tag{36}$$

Since $\bar{\sigma} > 0$, from (36) and (**??**) 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 (36) is true then we must have $\mu_W^* \leq \mu_S^*$. Moreover the following must also hold:

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

That is, we obtain $\mu_W^* > \mu_S^*$ which contradicts $\mu_W^* \leq \mu_S^*$. Therefore (36) cannot be true and we must have

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

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. Fix an allocation within a sound bank $c_S \equiv (\rho_S,\, c_{1S},\, c_{2S})$ such that budget constraint in (11) and the implementability constraint in (12) are both jointly satisfied.

First, suppose that the allocation in c_S has the following properties:

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

Then the contract:

$$c_1^k\left(0,\, \rho^k\right) = c_{1S} \text{ for } \rho^k = \pi \text{ and } c_1^k\left(0,\, \rho^k\right) = 0 \text{ for all } \rho^k > \pi$$

implements the allocation c_S as the unique equilibrium of the withdrawal game for the investors within the bank when the later is sound. To see that, suppose the bank is sound and a fraction $\epsilon \in [0,\, 1-\pi]$ of the patient investors within the bank request early withdrawal. Observe that any allocation where a positive fraction of the patient request early withdrawal cannot be an equilibrium when the bank is sound. Indeed, if $\epsilon > 0$ then the overall request for early withdrawals will be $\rho^k = \pi + \epsilon(1-\pi) > \pi$ and therefore the bank's contract would set $c_1^k = 0$, but then the budget constraint in (11) implies that requesting late withdrawal will result in a strictly larger payment for each patient investor. Therefore the only allocation consistent with equilibrium when the bank is sound corresponds to $\epsilon = 0$, that is, only the impatient request early withdrawals.

Second, suppose that the allocation in c_S has the following properties:

$$\rho_S > \pi, \; c_{1S} \geq c_{2S} \text{ and } \rho_S c_{1S} + (1-\rho_S)\frac{c_{2S}}{R} = 1-\tau$$

Consider a contract such that $c_1^k\left(0,\, \rho^k\right) > c_{1S}$ for $\rho^k < \pi$, $c_1^k\left(0,\, \rho^k\right) = c_{1S}$ for $\rho^k = \rho_S$ and $c_1^k\left(0,\, \rho^k\right) > c_{2S}$ for $\rho^k > \rho_S$. Then an argument similar to the one above can establish that conditional on being sound, the equilibrium of the withdrawal game is unique and implement the allocation in c_S. □

Proposition 4.

Proof. Suppose the banking regulator allows each bank to set its payment to c_{1S}^* as was assumed to be the case in the baseline economy analyzed in sections 4 and 5. The efficient allocation within sound banks was derived in section 3 and we denote it here by $c_S^* \equiv \left(\rho_S^*, c_{1S}^*, c_{2S}^*\right)$. Note that c_S^* is such that $\rho_S^* = \pi$ and $c_{1S}^* < c_{1S}^*$. In other words, c_S^* satisfies (11) and (12). Applying Propitiation 3 it follows that c_S^* can be implemented as an unique equilibrium of the withdrawal game within sound each bank. Moreover, for any other allocation such that $c_S \neq c_{1S}^*$ for which (11) - (12) is satisfied, it must be the case:

$$\begin{aligned} V_S\left(\rho_S^*,\, c_{1S}^*\right) &= \rho_S^* u\left(c_{1S}^*\right) + \left(1-\rho_S^*\right) u\left(c_{2S}^*\right) \\ &> \rho_S u\left(c_{1S}\right) + \left(1-\rho_S\right) u\left(c_{2S}\right) = V_S\left(\rho_S,\, c_{1S}\right) \end{aligned}$$

where the first and last equality above simply uses the definition of the objective function in (13) and the strict inequality on the second line follows from the fact that c_S^* maximizes the sum of investors expected utilities within sound banks as derived in Section 3.

Thus the sum of investors expected utilities within a given bank cannot be maximized unless the equilibrium allocation within the bank conditional on being sound is exactly c_S^*. Therefore an optimal contract must result in c_S^* as the unique equilibrium in all sound banks. One such contract was given in (14). $\square$

Proposition 5.

Proof. The allocation within a given weak bank can be summarized as follows: fraction ρ_W of its investors request to withdraw; the bank pays c_{1W} to the first θ and is then placed in resolution by the fiscal authority; each of the the remaining $(1-\theta)\hat{\rho}_W$ measure of impatient investors within this (and any other weak bank) receives $\hat{c}_1$ in period 1 and each of the remaining $(1-\theta)(1-\hat{\rho}_W)$ measure of patient investors within this (and any other weak bank) receives $\hat{c}_2$ in period 2 where, $\hat{c}_t$ for $t = 1, 2$, is determined by the economy-wide conditions in (21) - (22) and hence beyond the control of any given individual bank which would treat $\hat{c}_t$ as fixed.

First, if $\rho_W = \pi$ and $c_W \leq \hat{c}_2$ then (ρ_W, c_{1W}) satisfy the implementability condition for weak banks in (23) and the following contract:

$$c_1^k\left(\bar{\sigma}, \rho^k\right) = c_{1W} \text{ for } \rho^k = \pi \text{ and } c_1^k\left(\bar{\sigma}, \rho^k\right) = 0 \text{ for all } \rho^k > \pi$$

ensures that the equilibrium outcome within bank-k when weak is unique and corresponds to this allocation.

Second, if $\rho_W = 1$ and $c_W > \hat{c}_2$ then (ρ_W, c_{1W}) also satisfies (23). In this case a contract of the form:

$$c_1^k\left(\bar{\sigma}, \rho^k\right) = c_{1W} \text{ for all } \rho^k = 1$$

entails that the equilibrium outcome within bank-k when weak is unique and corresponds to the desired allocation.

Finally, the proof for the remaining combinations of ρ_W and c_{1W} for which the implementability constraint in (23) is satisfied (i.e. $\pi < \rho_W < 1$ and $c_{1W} = \hat{c}_2$) closely follows the above argument and is omitted. $\square$

Proposition 6.

Proof. First, if $c_1^k(\text{H}) = c_1^k(\text{H}\,|\text{NR})$ then we must have $c_1^k(\text{H}) \leq \hat{c}_2$. Otherwise, if $c_1^k(\text{H}) > \hat{c}_2$ then it also must be the case $c_1^k(\text{H}) > c_1^k(\text{H}\,|\text{NR})$ and therefore we obtain a contradiction. Given that $c_1^k(\text{H}) \leq \hat{c}_2$, any allocation such that $\rho_W = \pi$ and $c_{1W} \leq c_1^k(\text{H})$ satisfies the implementability constraints in (23) and, via Proposition 5, can be obtained as a unique equilibrium when bank-k is weak. Next, using the objective function in (24) and the relation in (25) we obtain that for any $c_{1W} < c_1^k(\mathrm{H})$:

$$V_W\left(c_1^k(\mathrm{H}),\, \pi\right) > V_W\left(c_{1W},\, \pi\right)$$

The above implies that whenever $c_1^k(\mathrm{H}) = c_1^k(\mathrm{H}\,|\mathrm{NR})$ it will be optimal for a weak bank to set $c_{1W} = c_1^k(\mathrm{H})$ and implement the allocation associated to $\left(\pi,\, c_1^k(\mathrm{H})\right)$. In other words, the optimal allocation within bank-k when weak will be characterized by $c_{1W} = c_1^k(\mathrm{H})$.

Second, if $c_1^k(\mathrm{H}) > c_1^k(\mathrm{H}\,|\mathrm{NR})$ and $c_1^k(\mathrm{H}\,|\mathrm{NR}) < \hat{c}_2$. Then it must be the case $c_1^k(\mathrm{H}) \leq \hat{c}_2 < c_1^k(\mathrm{H})$ and the implementability constraints in (23) entails that the bank can implement two classes of allocations: (i) $\rho_W = 1$ and $c_{1W} \in (\hat{c}_2,\, c_1^k(\mathrm{H})]$ and (ii) $\rho_W = \pi$ and $c_{1W} \leq c_1^k(\mathrm{H}\,|\mathrm{NR})$. Note that for any allocation with $\rho_W = 1$ the optimal is to set $c_{1W} = c_1^k(\mathrm{H})$ whereas for any allocation with $\rho_W = \pi$ the optimal is to set $c_{1W} = c_1^k(\mathrm{H}\,|\mathrm{NR})$. Hence, the optimal allocation within bank-k when weak will be characterized by $c_{1W} \in \left\{c_1^k(\mathrm{H}),\, c_1^k(\mathrm{H}\,|\mathrm{NR})\right\}$.

Third, assume $c_1^k(\mathrm{H}) > c_1^k(\mathrm{H}\,|\mathrm{NR})$ and $c_1^k(\mathrm{H}\,|\mathrm{NR}) = \hat{c}_2$ then in addition to the class of allocations in (i) and (ii) bank-k will be able to implement a third class of allocations: (iii) $\rho_W \in [\pi, 1]$ and $c_{1W} = c_1^k(\mathrm{H}\,|\mathrm{NR})$. In this case one can use (24) and (25) to show that for any $\rho_W > \pi$

$$V_W\left(c_1^k(\mathrm{H}\,|\mathrm{NR}),\, \pi\right) > V_W\left(c_1^k(\mathrm{H}\,|\mathrm{NR}),\, \rho_W\right)$$

Thus any allocations in (iii) is strictly inferior to the optimal among those in (ii), namely $\rho_W = \pi$ and $c_{1W} = c_1^k(\mathrm{H}\,|\mathrm{NR})$. Hence, the optimal allocation within bank-k when weak would still be characterized by $c_{1W} \in \left\{c_1^k(\mathrm{H}),\, c_1^k(\mathrm{H}\,|\mathrm{NR})\right\}$. □

Proposition 7.

Proof. In Proposition 6 we have shown that a weak bank would either set c_{1W} as high as possible, that is, $c_{1W} = c_1^k(\mathrm{H})$ or it will set c_{1W} as high as possible but still less or equal to $\hat{c}_2$, that is, $c_{1W} = c_1^k(\mathrm{H}\,|\mathrm{NR})$. First, assume $c_1^k(\mathrm{H}) \leq \hat{c}_2$ then we must have $c_1^k(\mathrm{H}) = c_1^k(\mathrm{H}\,|\mathrm{NR})$ and using that $\hat{c}_1 < \hat{c}_2$ (which follows from the first order condition in (19)) we obtain:

$$0 = u\left(c_1^k(\mathrm{H})\right) - u\left(c_1^k(\mathrm{H}\,|\mathrm{NR})\right) < (1-\pi)\left(u\left(\hat{c}_2\right) - u\left(\hat{c}_1\right)\right)$$

Moreover, in this case the optimal allocation clearly is $\rho_W = \pi$ and $c_{1W} = c_1^k(\mathrm{H}) = c_1^k(\mathrm{H}\,|\mathrm{NR})$. Second, if $c_1^k(\mathrm{H}) > c_1^k(\mathrm{H}\,|\mathrm{NR})$ then an optimal contract for bank-k must ensure the following:

$$\rho_W = \pi \text{ and } c_{1W} = c_1^k(\mathrm{H}\,|\mathrm{NR}) \text{ if } V_W\left(c_1^k(\mathrm{H}),\, 1\right) > V_W\left(c_1^k(\mathrm{H}\,|\mathrm{NR}),\, \pi\right)$$

$$\rho_W = 1 \text{ and } c_{1W} = c_1^k(\mathrm{H}) \text{ if } V_W\left(c_1^k(\mathrm{H}),\, 1\right) < V_W\left(c_1^k(\mathrm{H}\,|\mathrm{NR}),\, \pi\right)$$

Finally, if $V_W\left(c_1^k(\mathrm{H}),\, 1\right) = V_W\left(c_1^k(\mathrm{H}\,|\mathrm{NR}),\, \pi\right)$ then the bank will be indifferent between $\rho_W = \pi$ and $c_{1W} = c_1^k(\mathrm{H}\,|\mathrm{NR})$ and $\rho_W = 1$ and $c_{1W} = c_1^k(\mathrm{H})$ and therefore both of these allocations correspond to a contract which is optimal for the bank. □

Proposition 8.

Proof. First, suppose that the aggregate state is *good*. Then the regulator infers that all banks are sound, but still does not observe the requests for withdrawals within any of the banks. In this case, a regulatory allowing each bank to choose between 0 and c_{1S}^*. That is,

$$X_0 = \{0, c_{1S}^*\}$$

The above ensures that each bank finds it optimal to set up a contract which would uniquely results in the efficient within bank allocation

$$\rho^k = \pi,\ c_1^k = c_{1S}^* \text{ and } c_2^k = c_{12}^*$$

One such contract is the following: if the aggregate state is good, the bank sets $c_1^k(\rho^k, \sigma_k) = c_{1S}^*$ for $\rho^k = \pi$ and $c_1^k(\rho^k, \sigma_k) = 0$ for all $\rho^k > 0$. Finally, the level of the public good is equal to τ^*. In other words, the economy wide allocation will be efficient.

Second, suppose the aggregate state is *bad*. Since the regulator does not observe which are the sound and which are the weak banks the menu of payment options available to all banks must be the same X_n. Next, observe that a necessary condition for the equilibrium to be efficient in the bad aggregate state is that both c_{1W}^* and c_{1S}^* belong to the set of permitted payments. That is,

$$\{c_{1W}^*, c_{12}^*\} \subseteq X_n$$

So suppose all other weak banks select c_{1W}^* and consider a given weak bank choosing between c_{1W}^* and c_{1S}^*. A direct comparison of the expressions in (28) - (29) shows that the bank in question is able to deliver strictly greater utility to its investors by deviating and choosing c_{1S}^* and therefore setting c_{1W}^* is not incentive compatible. In other words, any economy-wide allocation where all weak banks select c_{1W}^* and, at the same time, c_{1S}^* is also an option is not consistent with equilibrium. Thus we have established the desired result, since the equilibrium in the bad aggregate state cannot be efficient unless both c_{1W}^* or c_{1S}^* are permitted by the regulator. □

Proposition 10.

Proof. We need to show that a bank run cannot occur when the fraction of weak banks in the bad aggregate state n is close enough to zero. First, the menu of payment options cannot be optimal unless all weak banks select a payment c_{1W} such that $c_{1W} \leq c^*_{1S}$ (see Proposition 11 below). Next, we show that for n sufficiently small, we must have

$$c_{1W} < \hat{c}_2 \tag{37}$$

Moreover, the above relation will hold regardless of whether weak banks experience a run or not. In other words, as the fraction of weak banks becomes converges to zero the bailout policy of the fiscal authority provides full deposit insurance and therefore weak banks will not be subject to runs from their investors. Observe that since $c^*_{1S} < c^*_{2S}$ in order to show (37) it will be sufficient to establish that as n converges to zero we have:

$$(\hat{c}_1, \hat{c}_2) \underset{n \to 0}{\longrightarrow} (c^*_{1S}, c^*_{2S}) \tag{38}$$

Indeed, holding other parameters fixed, as the fraction of weak banks in the bad aggregate state approaches zero we can use the conditions in (6) - (8) to obtain the following:

$$u'(c^*_{1S}) \underset{n \to 0}{\longrightarrow} v'(\tau^*)$$

$$u'(\hat{c}_1) = v'\left(\tau^* - n\hat{b}\right) \underset{n \to 0}{\longrightarrow} v'(\tau^*)$$

From the above it follows that regardless of whether the other weak banks experience a run, the payment plan for each bank in resolution will converge to the payment plan within sound banks as stated in (38) which, in turn, implies the desired condition in (37). ☐

Proposition 11.

Proof. The regulatory policy of the government is summarized in two sets X_0 and X_n where X_0 is the menu of payment options offered to each bank in the *good* aggregate state where all banks are sound and X_n is the menu of payment options offered to each bank in the *bad* aggregate state where fraction n of the banks are weak.

Next, let c^0_S be the within bank allocation in the good aggregate state, c^n_S the within bank allocation in the bad aggregate state when the bank is sound and c_W the within bank allocation in the bad aggregate state when the bank is weak. Note that we do not need to index c_W by n since there are no weak banks in the *good* aggregate state. Next, we define three sets:

$$\mathcal{F}_S(X_0) \equiv \left\{ c^0_S \;\middle|\; \text{such that } c^0_{1S} \in X_0, \text{ (11) and (12) holds.} \right\}$$

$$\mathcal{F}_{\mathcal{S}}\left(X_n\right) \equiv \left\{ c_S^n \;\middle|\; \text{such that } c_{1S}^n \in X_n, \text{ (11) and (12) holds.} \right\}$$

$$\mathcal{F}_{\mathcal{W}}\left(X_n,\, \hat{c}\right) \equiv \left\{ c_W \;\middle|\; \text{such that } c_{1W} \in X_n \text{ and (24) holds.} \right\}$$

Remark 1. If the following set of conditions is satisfied:

$$c_S^0 \in \mathcal{F}_{\mathcal{S}}\left(X_0\right),\; c_S^n \in \mathcal{F}_{\mathcal{S}}\left(X_n\right) \text{ and } c_W \in \mathcal{F}_{\mathcal{W}}\left(X_n,\, \hat{c}\right) \tag{39}$$

Then there is a banking contract such that:

(i) If the aggregate state is good, then the within bank allocation is unique and given by c_S^0.

(ii) If the aggregate state is bad and the bank is sound, then the within bank allocation is unique and given by c_S^n.

(iii) If the aggregate state is bad and the bank is weak, then the within bank allocation is unique and given by c_W.

Remark 1 follows directly from the Proposition 3 and Proposition 5. Next, for given menu of payment options in each aggregate state X_0 and X_n and for given resolution policy for the fiscal authority in the bad aggregate state $\hat{c} = (\hat{c}_1, \hat{c}_2)$ we can use Remark 1 to formulate the program of any given individual bank.

Remark 2. Taking as given the regulation and resolution policy of the government, X_0, X_n and $\hat{c}$, an individual bank will choose c_S^0, c_S^n and c_W in order to solve the following program:

$$\left\{ \begin{array}{c} \max\limits_{\left\{c_S^0, c_S^n, c_W\right\}} \left\{(1-q)V_S\left(c_S^0\right) + q\left[(1-n)V_S\left(c_S^n\right) + nV_W\left(c_W\right)\right]\right\} \\ \text{subject to (39)} \end{array} \right\} \tag{40}$$

If conditional on being weak, each bank chooses the same allocation then the bailout and resolution policy of the fiscal authority will be obtained as the solution to:

$$(1-\theta)\left[\hat{\rho}_W \hat{c}_1 + (1-\hat{\rho}_W)\frac{\hat{c}_2}{R}\right] = 1 - \tau - \bar{\sigma} - \theta c_{1W} + \hat{b} \tag{41}$$

$$u'\left(\hat{c}_1\right) = Ru'\left(\hat{c}_2\right) = v'\left(\tau - n\hat{b}\right) \tag{42}$$

Next, we want to formulate the program of the regulator. We set the tax rate equal to its value from the efficient allocation $\tau = \tau^*$. Allowing the regulator to choose τ will complicate the analysis without any new insights. Next, recall that any given bank chooses a contract so as to uniquely implement the most desirable allocation among these that can result in equilibrium of the withdrawal game within the bank taking as given the actions of the regulator, the fiscal authority and all other banks and investors. Now we pose a similar problem for the banking regulator.

Remark 3. The program of the banking regulator can be formulated as recommending an allocation to each bank c_S^0, c_S^n and c_W in addition to a bailout and resolution policy to the fiscal authority $\hat{c}$, $\hat{b}$ in order to maximize:

$$(1-q)\left[V_S\left(c_S^0\right)+v\left(\tau^*\right)\right]+q\left[(1-n)V_S\left(c_S^n\right)+nV_W\left(c_W^n\right)+v\left(\tau^*-n\hat{b}\right)\right]$$

Subject to (40), (41) and (42).

The current formulation of the program of the regulator is not complete, however, since we still need to specify the payment options offered to banks X_0 and X_n. We proceed by making the following observations.

Remark 4. An optimal regulatory policy in the good aggregate state is given by $X_0 = \{0, c_{1S}^*\}$.

Conditional on there being no weak banks, setting $X_0 = \{0, c_{1S}^*\}$ ensures that the efficient allocation is the unique equilibrium. Thus, if there are no weak banks, the regulatory options will be set to$\{0, c_{1S}^*\}$, the allocation is unique and corresponds to the efficient allocation derived in section 3. Henceforth we focus on the bad aggregate state and drop the superscript n when referring to allocations within sound banks.

Remark 5. It is without loss of generality if we suppose that in the bad aggregate state the banking regulator offers each bank a choice between 0, c_{1W} and c_{1S}.

Suppose that the payment options available to each bank are contained in the set X_n and consider an equilibrium outcome such that in the bad aggregate state all sound banks set c_{1S} and all weak banks set c_{1W}. Clearly it must be the case that c_{1W} and c_{1S} both belong to X_n. Moreover, if we were to remove all other element but these two from the set X_n then the equilibrium just described would still exist. The reason is that a weak bank that decided on c_{1W} among a set X containing, among c_{1W} and c_{1S}^n, also some other elements would still choose c_{1W} when these other elements were removed as payment options. The same argument holds for the sound banks as well.

Conversely, suppose we start from a smaller set where weak banks self-select into c_{1W} and sound banks self-select into c_{1S}. Notice that an addition of a new element, say z, to the set of options banks can choose from would either have no effect i.e. each bank would continue to make the same choice as before or the new element z would attract a subset of the banks that would abandon their old choice and switch to z. Suppose that adding z as an option entails weak banks choosing this payment (sound banks, in contrast, continue to prefer c_{1S} over both c_{1W} and z).

The situation just described, however, is equivalent to replacing c_{1W} with z. Indeed, the addition of z to the set of payments available to each bank makes c_{1W} an off-equilibrium choice and the resulting equilibrium allocation would also obtain if c_{1W} were not available in the first place. Finally, observe that this conclusion remains true regardless of whether the regulator actually prefers weak banks to choose z instead of c_{1W}. At the same time, if the regulator preferred z to c_{1W} for the weak banks then deleting c_{1W} rather than keeping it as an option cannot reduce welfare.

To summarize: the above argument suggest that it will be without loss of generality if we assume that each bank is offered at most three options: 0, c_{1W} and c_{1W} (which would reduce to just two whenever $c_{1W} = c_{1S}$). As pointed out previously, the option to set zero is necessary to prevent runs and to ensure that the equilibrium where weak banks self-select into c_{1W} whereas sound banks self-select into c_{1S} not only exists but is unique as well.

Next, consider the economy-wide allocation E_B in the bad aggregate state where in addition to c_{1W} and c_{1S} the remaining components of the allocation are collected below:

$$E_B \equiv \left(\rho_W, \rho_S, c_{1S}, c_{2S}, c_{1W}, \hat{c}_1, \hat{c}_2, \hat{b}\right)$$

Remark 6. If the aggregate state is bad, then the regulatory policy $X = \{0,\, c_{1W},\, c_{1S}\}$ can result in the economy-wide allocation E as an equilibrium if and only if E jointly satisfies (11) - (12), (32) - (33) and (19) - (22)

First, from Proposition 3 the allocation within sound banks will satisfy (11) - (12). Second, from Proposition 5 the allocation within weak banks will satisfy (32). Third, the ex-post optimal bailout and resolution policy of the fiscal authority will satisfy (19) - (22). Fourth, weak banks must prefer to implement the allocation associated with c_{1W} and sound banks must prefer to implement the allocation associated with c_{1S}. That is, the incentive compatibility constraints in (32) - (33) must hold.

Define $\mathcal{F}_{\mathcal{R}}$ to denote the set of all economy-wide allocations E_B for which (11) - (12), (32) - (32) - (33) and (19) - (22) holds. The optimal regulation policy in the bad aggregate state can be stated as choosing the economy-wide allocation to maximize the sum of investors' expected utilities in the bad aggregate state.

$$\max_{E_B \in F_R} \left\{(1-n)V_S\left(c_S^n\right) + nV_W\left(c_W^n\right) + v\left(\tau - n\hat{b}\right)\right\} \tag{43}$$

□