DISCUSSION PAPER SERIES

No. 4778

TOO MANY TO FAIL – AN ANALYSIS OF TIME INCONSISTENCY IN BANK CLOSURE POLICIES

Viral V Acharya and Tanju Yorulmazer

FINANCIAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

TOO MANY TO FAIL – AN ANALYSIS OF TIME INCONSISTENCY IN BANK CLOSURE POLICIES

Viral V Acharya, London Business School (LBS) and CEPR
Tanju Yorulmazer, New York University

ABSTRACT

Too Many to Fail – An Analysis of Time Inconsistency in Bank Closure Policies*

This Paper shows that bank closure policies suffer from a 'too many to fail' problem: when the number of bank failures is large, the regulator finds it *ex post* optimal to bail out some or all failed banks, whereas when the number of bank failures is small, failed banks can be acquired by the surviving banks. This gives banks incentives to herd and increases systemic risk, the risk that many banks may fail together. The *ex post* optimal regulation may thus be sub-optimal from an *ex ante* standpoint. We formalize this time-inconsistency of bank regulation. We also argue that by allowing banks to purchase failed banks at discounted prices and by partially nationalizing the bailed-out banks, a regulator may be able to mitigate the induced systemic risk.

Viral V Acharya
Institute of Finance and Accounting
London Business School
6 Sussex Place, Regent's Park
London
NW1 4SA
Tel: (44 20) 7262 5050 x 3535
Fax: (44 20) 7724 3317
Email: vacharya@london.edu

Tanju Yorulmazer
Financial Industry and Regulation
Division
Bank of England
Threadneedle Street
London
EC2R 8AH
Tel: (44 20) 7601 3198
Fax: (44 20) 7601 3217
Email: tanju.yorulmazer@bankofengland.co.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=147901

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=158591

1 Introduction

Historically, central banks evolved as a response to wide-spread banking crises (Gorton, 1985). The central banks were seen as crises managers who may rescue troubled banks in times when closures could exacerbate welfare losses.[1] Over time, central banks have also taken on the role of crises prevention and often justified prudential regulation norms as a way of mitigating systemic risk, the risk that many banks may fail together. In this paper, we argue that the crises-prevention role of a central bank (more generally, the bank regulator) conflicts with its crises-management role due to a lack of commitment in optimal policies. This lack of commitment induces bank behavior that increases the likelihood of systemic banking crises.

Specifically, time-consistent (or subgame-perfect) regulation of banks suffers from a "too-many-to-fail" problem: In order to avoid continuation losses, a regulator finds it ex-post optimal to bail out banks when the number of failures is large; in contrast, if only some banks fail, then these banks can be acquired by the surviving banks. In particular, as the number of failed banks increases and the number of surviving banks decreases, the investment opportunity set for surviving banks becomes larger but the total investment capacity of surviving banks decreases. Thus, it becomes more likely that some banks would have to be liquidated to investors outside the banking sector resulting in a loss of continuation values. In turn, it becomes optimal for the regulator to bail out some of these failed banks instead of liquidating them.[2]

This too-many-to-fail problem induces banks to herd ex ante in order to increase the likelihood of being bailed out. For example, they may lend to similar industries or bet on common risks such as interest and mortgage rates. This leads to too many systemic banking crises. The central bank's problem is thus one of time-inconsistency. Its ex-post optimal bailout policy is not ex-ante optimal. Or said differently, the ex-ante optimal policy would involve not rescuing banks in crises, but this is not time-consistent. We formalize these ideas in a framework wherein the ex-ante and the ex-post optimal policies are endogenously derived

---

[1]See, for example, Hoggarth, Reis and Saporta (2001) who provide evidence on the real costs of systemic banking crises over the world as whopping 15–20% of annual GDP based on measures of cumulative GDP gaps during and following the crises.

[2]Hoggarth, Reidhill and Sinclair (2004) study resolution policies adopted in 33 systemic crises over the world during 1977–2002. They document that when faced with individual bank failures authorities have usually sought a private-sector resolution in which losses have been passed onto existing shareholders, managers, and sometimes uninsured creditors, but not to taxpayers. However, government involvement has been an important feature of the resolution process during systemic crises: At early stages, liquidity support from central banks and blanket government guarantees have been granted, usually at a cost to the budget; bank liquidations have been very rare and creditors have rarely made losses; finally, bank mergers have been employed only at the restructuring phases. Kasa and Spiegel (1999) and Santomero and Hoffman (1999) also provide evidence that regulatory actions appear to depend on whether the problems are idiosyncratic, that is, specific to particular institutions, or systemic with potential threats to the whole system.

based on a well-specified objective function for the regulator.

To make our main point, we consider the simplest possible setting. In particular, we consider a two-period, two-industry model with two banks, a regulator, and outside investors who can purchase banking assets were they to be liquidated. Two central assumptions drive our results: (i) banks are more efficient users of banking assets than outsiders as long as they take good projects,[3] and (ii) there is a possibility of moral hazard in that bank owners derive private benefits from bad projects; hence, banks take good projects only if bank owners are given a large enough share in bank profits. We require that each bank invests in one of the two industries (we relax this assumption in an addendum to the paper). Banks choose whether to invest in the same industry or in different ones. This decision affects the correlation of bank returns and in turn the likelihood that banks will fail together. The regulator designs closure and bailout policies in order to maximize the total output generated by the banking sector net of any costs associated with closures and bailouts. These policies are assumed to be rationally anticipated by banks and depositors.

If the bank return from the first-period investment is high, then the bank operates one more period and makes the second-period investment. If the bank return from the first-period investment is low, then there is a run on the bank. If there is a surviving bank, then we assume that it has resources from its first-period profits to purchase the failed bank (an assumption we relax in the n-bank case contained in a companion paper). The regulator decides whether to intervene or to let the surviving bank (if any) purchase the failed bank. When a regulator intervenes, the decision is whether to keep the bank open through a bailout or to liquidate its assets to outside investors. When a bank is bailed out, the regulator pays off the deposits that are in default. In return for this injection of funds, the regulator may dilute the equity share of bank owners. The immediacy of funds provided to the bank entails fiscal costs for the regulator (assumed to be exogenous to the model).

Assumption (i) generates an allocation inefficiency from liquidating assets to outsiders. Since bailouts are also costly, the regulator does not intervene when only one bank fails; assets of the failed bank are sold to the surviving bank, the efficient user of these assets. The surviving bank captures a surplus from its superior skills in running banking assets compared to outsiders, but there are no welfare losses. However, when both banks fail, one or both banks may be bailed out if the costs of injecting funds are smaller than the efficiency costs of liquidating assets to outsiders. This gives rise to a "too-many-to-fail" problem. Crucially, unlike the state in which a single bank fails, the joint-failure state always entails some welfare losses.

---

[3]James (1991) studies the losses from bank failures in the United States during the period 1985 through mid-year 1988, and documents that "there is significant going concern value that is preserved if the failed bank is sold to another bank (a "live bank" transaction) but is lost if the failed bank is liquidated by the Federal Deposit Insurance Corporation (FDIC)."

Ex ante, the regulator wishes to implement a low correlation between banks' investments in order to minimize the likelihood of the joint-failure state, and simultaneously implement closure policies that are ex-post optimal. The regulator can implement such a welfare-maximizing outcome if it can commit to sufficiently diluting the share of bank owners in bailed-out banks when banks have failed together. With sufficient dilution, the bailout subsidies are small, and ex ante banks invest in different industries to capture the gains from purchasing the failed bank when they survive. However, assumption (ii) implies that such a dilution may not always be feasible. If the moral hazard due to private benefits is sufficiently high, then excessive dilution of a bank's equity leads bank owners to choose bad projects and this generates continuation values that are worse than liquidation values. In this case, the only credible mechanism through which the central bank can implement low correlation is committing to liquidate banks in the joint-failure state. In general, this is ex-post inefficient and thus lacks commitment. In turn, this lack of commitment gives rise to an incentive among banks to invest in the same industry in order to capture bailout subsidies in joint-failure states.

In a companion paper, we check the robustness of our results in an extension with n banks where each bank invests either in a common asset or in a bank-specific asset. In particular, even though each surviving bank has enough resources to buy at least one failed bank, the total resources of surviving banks may not suffice for purchasing all failed banks when the number of bank failures is large. On the one hand, the surplus captured by surviving banks is larger when only a few banks survive as there is less competition for purchasing failed banks. On the other hand, the regulator finds it ex-post optimal to bail out more of these failed banks since not all of them can be paid for by resources of the surviving banks.

Finally, we check that our herding results are not an artifact of the assumption that banks invest all funds in just one asset. We consider in the addendum to this paper a second extension with two banks where each bank chooses what shares of its portfolio to invest in a common asset and a bank-specific asset. Thus, banks effectively choose from a complete spectrum of inter-bank correlations. We show that the intuition from the benchmark model prevails in both these extensions. When the moral hazard due to private benefits is sufficiently severe, bailouts do not involve sufficient dilution of bank owners' equity. In turn, the anticipation of bailout subsidies in joint-failure states induces banks to herd by investing excessively in the common asset.

From a policy perspective, the main contribution of our paper is to provide a framework for jointly analyzing the central bank policies that are aimed at resolution and prevention of systemic crises. Though systemic risk is the main rationale provided for justifying prudential regulation norms, the focus of these norms is almost exclusively at the level of an individual bank. Our paper argues that the genesis of inefficient systemic risk may potentially lie in the very crises-management role of central banks. Thus, the paper highlights the need for

understanding and designing regulatory policies at a systemic level rather than only at an individual bank level. For example, our analysis implies that if the regulator can transfer additional surplus to surviving banks by allowing them to purchase failed banks at discounts relative to the prices that outside investors are willing to pay (that is, at opportunity costs to depositors of the failed banks), then bank incentives to herd and the resulting systemic risk can be mitigated.

The idea that bank regulation may be time-inconsistent and may induce moral hazard is not new, but our specific application of these ideas is novel. For example, Mailath and Mester (1994) and Freixas (1999) discuss the time-inconsistency of closure policy in a single-bank model, and Bagehot (1873) in his famous piece on bank bailouts discusses the moral hazard from rescuing failed banks. The focus in the existing literature is however on an individual bank and its choice of risk rather than on multiple banks and their choice of joint-failure risk.

In a deviation from the standard literature, Penati and Protopapadakis (1988) assume that the regulator provides insurance to uninsured depositors when the number of banking failures is large, and illustrate that this leads banks to invest inefficiently in common markets (as in our second extension) so as to attract deposits at a cheaper cost. In contrast, we endogenize the ex-post bailout policies of the regulator. Furthermore, deposit rate in our model is set before banks make investment decisions. Hence, banks herd in our model not because if affects deposit rates but in order to capture the (endogenously derived) bailout subsidies. Mitchell (1997) considers an argument along the lines of the "signal-jamming" model of Rajan (1994) to show that if the regulator bails out banks when they fail together, then banks coordinate on disclosing their losses and delay classifying bad loans by rolling them over. Again, the too-many-to-fail problem is exogenous to her model since the regulator's objective as well as closure policies are not explicitly modelled.

Perotti and Suarez (2002) consider a dynamic model where selling failed banks to surviving banks (reducing competition) increases the charter-value of surviving banks and gives banks ex-ante incentives to stay solvent. This strategic benefit is present in our model in a different guise as the discount at which surviving banks purchase failed banks. However, in contrast to our model, their paper does not examine the effect of closure policies on inter-bank correlation. Diamond and Rajan (2003) show that a bank failure can cause aggregate liquidity shortages and regulatory intervention may be optimal. The focus of their paper is on demonstrating the liquidity-related channel of contagion and the difficulty of resolving it ex post, but not on its implications for the ex-ante investment choices of banks. Finally, some of the ideas presented in this paper are motivated by the analysis in Acharya (2001).

The remainder of the paper is structured as follows. Section 2 and Section 3 present the model and the analysis. Section 4 concludes and the proofs that are not in the text are contained in the Appendix. For sake of completeness, we have included as an Addendum the

somewhat lengthy analysis of the second extension. We do not envisage this as being a part of the final version of the paper.

2   Benchmark model

The model is outlined in Figure 1. We consider an economy with three dates – $t = 0, 1, 2$, two banks – Bank A and Bank B, bank owners, depositors, outside investors, and a regulator. Each bank can borrow from a continuum of depositors of measure 1. Bank owners as well as depositors are risk-neutral, and obtain a time-additive utility w_t where w_t is the expected wealth at time t. Depositors receive a unit of endowment at $t = 0$ and $t = 1$. Depositors also have access to a reservation investment opportunity that gives them a utility of 1 per unit of investment. In each period, that is at date $t = 0$ and $t = 1$, depositors choose to invest their good in this reservation opportunity or in their bank.

Deposits take the form of a simple debt contract with maturity of one period. In particular, the promised deposit rate is not contingent on investment decisions of the bank or on realized returns. Finally, the dispersed nature of depositors is assumed to lead to a collective-action problem, resulting in a run on a bank that fails to pay the promised return to its depositors. In other words, the deposit contract is "hard" and cannot be renegotiated. In order to keep the model simple and yet capture the fact that there are limits to equity financing due to associated costs (for example, due to asymmetric information as in Myers and Majluf, 1984), we do not consider any bank financing other than deposits.

Banks require one unit of wealth to invest in a risky technology. The risky technology is to be thought of as a portfolio of loans to firms in the corporate sector. The performance of the corporate sector determines its random output at date $t + 1$. We assume that all firms in the sector can either repay fully the borrowed bank loans or they default on these loans. In case of a default, we assume for simplicity that there is no repayment.

Suppose R is the promised return on a bank loan. We denote the random repayment on this loan as $\widetilde{R}$, $\widetilde{R} \in \{0, R\}$. The probability that the return from these loans is high in period t is α_t:

$$\widetilde{R} = \begin{cases} R & \text{with probability} \quad \alpha_t, \\ 0 & \text{with probability} \quad 1 - \alpha_t. \end{cases} \tag{1}$$

We assume that the returns in the two periods are independent but allow the probability of high return to be different in the two periods. This helps isolate the effect of each probability on our results.

There is a potential for moral hazard at the level of an individual bank. If the bank chooses a bad project, then when the return is high, it cannot generate R but only $(R - \overline{\Delta})$

6

and its owners enjoy a non-pecuniary benefit of $B < \overline{\Delta}$. Therefore, for the bank owners to choose the good project, appropriate incentives have to be provided by giving bank owners a minimum share of the bank's profits. We denote the share of bank owners as θ. If r is the cost of borrowing deposits, then the incentive-compatibility constraint is:

$$\alpha_t \theta (R - r) \geqslant \alpha_t \left[\theta((R - \overline{\Delta}) - r) + B \right]. \qquad (IC) \tag{2}$$

We have assumed that the bank is able to pay the promised return of r when the investment had the high return irrespective of whether the project is good or bad. The left hand side of the (IC) constraint is the expected profit for the bank from the good project when it has a share of θ of the profit. On the right hand side, we have the expected profit from the bad project when bank owners have a share of θ, plus the non-pecuniary benefit of choosing the bad project. Using this constraint, we can show that bank owners need a minimum share of $\overline{\theta} = \frac{B}{\overline{\Delta}}$ to choose the good project.[4] We assume that at $t = 0$, the entire share of the bank profits belongs to the bank owners, and therefore, there is no moral hazard to start with.

In addition to banks and depositors, there are outside investors who always have funds to purchase banking assets were these assets to be liquidated. However, outsiders do not have the skills to generate the full value from banking assets. In particular, outsiders are inefficient users of banking assets relative to the bank owners provided bank owners operate good projects. to capture this, we assume that outsiders cannot generate R in the high state but only $(R - \Delta)$. Thus, when the banking assets are liquidated to outsiders, there may be a social welfare loss due to misallocation of these assets. We also assume that $\overline{\Delta} > \Delta$ so that outside users of the banking assets can generate more than what the banks can generate from the bad project.

Finally, there is a regulator whose objective is to choose closure and bailout policies in order to maximize the total output generated by the banking sector net of any costs associated with closures and bailouts. These policies are assumed to be rationally anticipated by banks and depositors. Below we describe these policies informally. The formal description follows in the model analysis.

If the bank return from the first-period investment is high, then the bank operates one more period and makes the second-period investment. If the bank return from the first-period investment is low, then there is a run on the bank. We assume that if there is a surviving bank, then it has resources from its first-period profits to purchase the failed bank. The regulator decides whether to intervene or to let the surviving bank (if any) purchase the failed bank. When a regulator intervenes, the decision is whether to keep the bank open through a bailout or to liquidate its assets to outside investors. When a bank is bailed out,

---

[4]See Hart and Moore (1994) and Holmstrom and Tirole (1998) for models with similar incentive-compatibility constraints.

Bank A\Bank B	High	Low
High	SS	SF
Low	FS	FF

Table 1: Possible outcomes from first period investments.

$A \setminus B$	High	Low
High	α_t	0
Low	0	$1 - \alpha_t$

Table 2: Joint distribution of bank returns when they invest in the same industry.

the regulator pays off the deposits that are in default. In return for this injection of funds, the regulator may dilute the equity share of bank owners. The immediacy of funds provided to the bank entails fiscal costs for the regulator (assumed to be exogenous to the model). These costs are also a part of the regulator's objective function.

Therefore, the possible states at date 1 are given as follows, where S indicates survival and F indicates failure:

SS : Both banks had the high return, and they operate in the second period.

SF : Bank A had the high return, while Bank B had the low return. Bank B is bailed out, or acquired by bank A, or liquidated.

FS : This is the symmetric version of state SF.

FF : Both banks failed. Banks are either bailed out or liquidated (one or both).

2.1   Correlation of bank returns

A crucial aspect of our model is the correlation of bank returns. At date 0, banks borrow deposits and then they choose the composition of loans that compose their respective portfolios. This choice determines the level of correlation between the returns from their respective investments. We refer to this correlation as "inter-bank correlation".

We suppose that there are two possible industries in which banks can invest, denoted as 1 and 2. Bank A (B) can lend to firms A_1 and A_2 (B_1 and B_2) in industries 1 and 2,

$A \setminus B$	High	Low
High	α_t^2	$\alpha_t(1 - \alpha_t)$
Low	$\alpha_t(1 - \alpha_t)$	$(1 - \alpha_t)^2$

Table 3: Joint distribution of bank returns when they invest in different industries.

8

respectively. If in equilibrium banks choose to lend to firms in the same industry, specifically they either lend to A_1 and B_1, or they lend to A_2 and B_2, then they are assumed to be perfectly correlated. However, if they choose different industries, then their returns are less than perfectly correlated, say independent. When banks invest in the same industry, the correlation of banks' returns is $\rho = 1$, whereas, when they invest in different industries, we have $\rho = 0$. This gives us the joint distribution of bank returns as given in Tables 2 and 3. Note that the individual probability of each bank succeeding or failing is constant (α_0 and $1 - \alpha_0$, respectively).

Inter-bank correlation ρ is a *joint* choice of the banks which could be interpreted as the outcome of a cooperative game between banks. In our model, this joint choice of inter-bank correlation is identical to the one that arises from a non-cooperative (Nash) equilibrium choice of industries by banks playing a coordination game.

3   Analysis

We analyze the model proceeding backwards from the second period to the first period. We examine separately the outcomes under each of the four states at $t = 1$: SS, SF, FS, and FF. The promised deposit rate in state i at $t = 1$ is denoted as r_1^i, and in the first period as r_0. We assume throughout that $R > r_0$ and $R > r_1^i$.

3.1   Both banks survived (SS):

In this case, both banks operate for another period. Since, the returns from each period's investments are assumed to be independent, the probability of having the high return for each bank is equal to α_1. This is the last period and there is no further investment opportunity. Thus, depositors get the promised rate r_1^{ss} when the bank has the high return and they get 0 when the return is low. In turn, the promised rate that satisfies the individual rationality of depositors is $r_1^{ss} = 1/\alpha_1$. The expected payoff to the bank from its second-period investment when both banks survived, $E(\pi_2^{ss})$, is thus

$$E(\pi_2^{ss}) = \alpha_1[R - r_1^{ss}] = \alpha_1 R - 1. \tag{3}$$

Note that this payoff is independent of inter-bank correlation.

3.2   Only one bank survived (SF or FS):

This is the case where one bank had the high return while the other had the low return. Note that this state has a positive probability only when banks invest in different industries.

Without loss of generality, we concentrate on the case SF where Bank A had the high return and Bank B had the low return.

For a bank to continue operating for another period, it needs to pay its old depositors r_0 and it needs an additional one unit of wealth for the second investment. The failed bank B cannot generate the needed funds, $(1 + r_0)$, from its depositors at $t = 1$: Its depositors are endowed with only one unit of wealth at $t = 1$. Anticipating this, depositors run on the bank and the bank fails. An important possibility is that the surviving bank A may be able to purchase the assets of the failed bank B. Next, we argue that it is optimal for the regulator not to intervene and let Bank A purchase Bank B's assets. Indeed, we also show that it is optimal for Bank A to do so.

Profitability of Asset Purchases   The sale of Bank B's assets to Bank A may result in a transfer of wealth between banks, but it is not associated with any welfare loss as long as assets stay within the banking system and banks choose good projects. In contrast, the sale of Bank B's assets to outsiders is a source of social welfare loss since Bank A is a more efficient user of these assets. Therefore, as long as Bank A acquires Bank B's assets, the regulator does not have an incentive to intervene ex-post: in the continuation game, the regulator does not care about which bank owns these assets.

To show that it is indeed profitable for the surviving bank to acquire failed bank's assets, we make the following set of assumptions:

(i) Bank A makes a take-it-or-leave-it offer to Bank B: Note that outsiders can generate a maximum return of $(R - \Delta)$ in the high-return state from assets of Bank B. However, there is a refinancing cost of one unit for these assets to be reinvested. Therefore, outsiders are willing to pay a maximum of $[\alpha_1(R - \Delta) - 1]$ for these assets. If Bank A has all the bargaining power, then because of its special skills it can purchase the assets of Bank B at a "discount": Bank A can make a take-it-or-leave-it offer that is equal to $[\alpha_1(R - \Delta) - 1]$, in other words, equal to the amount the outsiders are willing to pay. The main thrust of our results would not change if the surviving bank is assumed to have a partial bargaining power, as long as it can capture some of the surplus arising from its efficient management of assets relative to the outsiders.

(ii) Bank A has enough funds to purchase assets of Bank B: We assume that the necessary funds for purchasing the assets come from the first period profits. We assume that R is high enough so that the first-period profits are sufficient for the asset purchase. In particular, the condition is $(R - r_0) \geqslant [\alpha_1(R - \Delta) - 1]$, which can be simplified to the condition $R \geqslant \frac{(r_0 - \alpha_1\Delta - 1)}{(1 - \alpha_1)}$.

(iii) Bank A has access to depositors of Bank B only after the purchase: The acquiring bank needs one unit of wealth for its own investments, one unit for the purchased

10

bank's investments, and $[\alpha_1(R - \Delta) - 1]$ units to purchase failed bank's assets, therefore $[1 + \alpha_1(R - \Delta)]$ in total. We assume that acquiring the assets of Bank B also enables Bank A to access Bank B's depositors (operate Bank B's "branches"). Thus, Bank A can borrow one unit from its own depositors and one unit from Bank B's depositors for second period investments.

(iv) Depositors of Bank B share equally the proceeds from the asset sale: We assume that the proceeds from the sale of Bank B's assets, $[\alpha_1(R - \Delta) - 1]$, are equally distributed among its depositors. We make the simplifying assumption that $[\alpha_1(R - \Delta) - 1] < 1$ as otherwise deposits can always be paid back fully and would be riskless.

Suppose now that the existing assets of Bank A and the acquired assets of Bank B are "merged." The value of Bank A's portfolio depends on the correlation of these assets. There are four possibilities for the return from the pooled portfolio: (R, R), $(R, 0)$, $(0, R)$, $(0, 0)$. Two cases arise: In the first case, R is high enough so that Bank A can honor promises to its depositors whenever one of the assets performs well; in the second case, R is low and both assets need to perform well to enable Bank A to make promised payments to depositors.

In the first case, the four possible states at $t = 2$ and Bank A's payoffs in these states are:

$$\pi_2^{sf} = \begin{cases} 2R - 2r_1^{sf} & if \quad (R, R) \\ R - 2r_1^{sf} & if \quad (R, 0) \text{ or } (0, R) \\ 0 & if \quad (0, 0) \end{cases} . \tag{4}$$

In the second case, the four possible states and Bank A's payoffs are:

$$\pi_2^{sf} = \begin{cases} 2R - 2r_1^{sf} & if \quad (R, R) \\ 0 & otherwise \end{cases} . \tag{5}$$

From the individual rationality of depositors, we can calculate $r_1^{sf} = 1/\alpha_1^{sf}$, where α_1^{sf} is the probability that the depositors get their promised return.

Now the question is: will Bank A be willing to buy Bank B's assets? Note that the expected borrowing cost for a bank will always be 1 due to the risk-neutrality of depositors. However, Bank A buys Bank B's assets at a discount equal to $\alpha_1\Delta$, the surplus from its efficient management of assets relative to the outsiders. Therefore, Bank A is always better off purchasing Bank B's assets.

Lemma 1 *It is profitable for Bank A to purchase assets of Bank B. The value of Bank A increases by $\alpha_1\Delta$ from the asset purchase.*

Proof: *For brevity, we give the proof only for the first case where Bank A can pay off depositors if either of the assets perform well. The proof for the second case is analogous and*

11

is provided in the Appendix. In the first case, $\alpha_1^{sf} = 1 - \Pr(0,0) = \alpha_1(2 - \alpha_1)$. Let $\pi_2^{sf}(0)$ be the bank's profit in state SF when banks invest in different industries, that is, $\rho = 0$. Note that if banks invest in the same industry then state SF does not arise. Using the probability distribution in equation (4) the expected profit of Bank A from the second-period investment can be written as:

$$E(\pi_2^{sf}(0)) = \Pr(R,R)\ (2R - 2r_1^{sf}) + 2\Pr(R,0)\ (R - 2r_1^{sf}) - [\alpha_1(R - \Delta) - 1]. \tag{6}$$

Using $r_1^{sf} = \frac{1}{\alpha_1^{sf}} = \frac{1}{\alpha_1(2-\alpha_1)}$, we get

$$
\begin{aligned}
E(\pi_2^{sf}(0)) &= 2R\ \underbrace{[\Pr(R,R) + \Pr(R,0)]}_{=\alpha_1} - 2r_1^{sf}\ \underbrace{[\Pr(R,R) + 2\Pr(R,0)]}_{=\alpha_1^{sf}} - [\alpha_1(R - \Delta) - 1] \\
&= 2\alpha_1 R - 2 - [\alpha_1(R - \Delta) - 1] = \alpha_1 R - 1 + \alpha_1 \Delta. \tag{7}
\end{aligned}
$$

Note that if the bank does not purchase the assets, its expected profit is equal to $(\alpha_1 R - 1)$. By purchasing the assets the bank increases its expected profits by $\alpha_1 \Delta > 0.\diamondsuit$

3.3 Both banks failed (FF)

Since there is no social welfare loss when assets stay within the banking system, the regulator does not have any incentive to intervene in states SS and SF (and FS). However, in state FF, assets of failed banks can be purchased only by outside investors who are not the most efficient users of these assets. Hence, the regulator compares the welfare loss resulting from asset sales to outsiders with the cost of bailing out one or both of the failed banks. If it turns out that the welfare loss from liquidation is greater, then the regulator may decide to intervene. The regulator's ex-post decision is thus more involved in state FF and we examine it fully.

Note that for a bank to continue operating, it needs r_0 units to pay old deposits and an additional one unit to undertake the second period investment, a total of $(r_0 + 1)$ units. Since available deposits for a bank amount to only one unit (the $t = 1$ endowment of its depositors), the bank cannot operate unless the regulator injects r_0 at $t = 1$. In order to analyze the regulator's decision to bail out or close one or both of the banks, we make the following assumptions:

(i) The regulator incurs a cost of $f(x)$ when it injects x units of funds into the banking sector. We assume this cost function is increasing and convex: $f' > 0$ and $f'' > 0$. We do not model this cost for which we have in mind fiscal and opportunity costs to the regulator from providing funds with immediacy to the banking sector. Thus, if the regulator bails out only one bank, it incurs a bail out cost of $f(r_0)$. If the regulator bails out both failed banks, then the cost is $f(2r_0) > 2f(r_0)$.

(ii) The regulator collects an amount of tax τ from depositors of the bailed-out bank. In other words, the net subsidy to depositors of this bank equals $(r_0 - \tau)$.

(iii) The net subsidy received by depositors of a bailed-out bank equals their payoff were the bank to be closed. This implies that when one bank is closed and the other bailed out, depositors of both banks receive the same payoff. Note that if a bank is closed, that is, its assets are liquidated to outsiders, then depositors of the bank receive $[\alpha_1(R - \Delta) - 1] < 1$. This is the amount the outsiders are willing to pay for the assets (as in state SF). We assume that the tax rate τ is such that the net subsidy to depositors of a bailed-out bank, $(r_0 - \tau)$, equals $[\alpha_1(R - \Delta) - 1]$. In other words, $\tau = [r_0 - \alpha_1(R - \Delta) + 1]$.[5]

(iv) The regulator can take an equity share in the bailed out bank(s). Let β be the share the regulator takes in a bailed out bank. If the bailed out bank has a high return from the second investment (which is likely with a probability α_1), then the regulator gets back $\beta(R - r_1^{ff})$ at $t = 2$. Ex post, such dilution of bailed-out bank's equity is merely a transfer from bank owners to the regulator. However, as argued before, if the regulator takes a share greater than $(1 - \bar{\theta})$, then the bank owners are left with a share of less than $\bar{\theta}$, the critical share below which the bank chooses the bad project. Since liquidating the bank generates a higher payoff compared to that from a bailed-out bank that chooses a bad project ($\overline{\Delta} > \Delta$), the regulator never takes a share greater than $(1 - \bar{\theta})$.

Under these assumptions, the regulator's bail out policy can be characterized as follows. The regulator's objective in state FF is to maximize the total expected output of the banking sector net of any bailout or closure costs. We denote this as $E(\Pi_2^{ff})$. Thus, if both banks are closed, the regulator's objective function takes the value

$$E(\Pi_2^{ff}) = 2\left[\alpha_1(R - \Delta) - 1\right], \tag{8}$$

which is the liquidation value of banking assets to outsiders. This equals $[2(\alpha_1 R - 1) - 2\alpha_1 \Delta]$, the difference between the banking sector output in each of the states SS, SF, and FS, minus the liquidation costs from closing both banks.

If one bank is bailed out and the other closed, then the regulator's objective function takes the value

$$E(\Pi_2^{ff}) = (\alpha_1 R - 1) - f(r_0) + [\alpha_1(R - \Delta) - 1], \tag{9}$$

where the first two terms represent the value of bailed-out bank net of bailout costs, and the third term represents the liquidation value realized from closure of the other bank. This can

---

[5]We stress that this is only a simplifying assumption which ensures that the borrowing rate of banks at $t = 0$, r_0, does not depend on inter-bank correlation. If the net subsidy to depositors is greater in state FF relative to state SF, then r_0 would be lower for high inter-bank correlation and the banks would have stronger incentives to invest in the same industry. This would make our results on ex-ante herding by banks only stronger.

be expressed in the intuitive form $[2(\alpha_1 R - 1) - \alpha_1 \Delta - f(r_0)]$, the difference between the banking sector output in each of the states SS, SF, and FS, minus the liquidation cost from closing one bank and the bailout costs from rescuing the other.

Finally, if both banks are bailed out, then the regulator's objective function takes the value

$$E(\Pi_2^{ff}) = 2(\alpha_1 R - 1) - f(2r_0), \tag{10}$$

as the bailout costs are now based on the total amount of funds, $2r_0$, injected into the banking sector with immediacy.

Comparing these objective-function evaluations, we obtain the following closure/bailout policy for the regulator in state FF. It has the intuitive property that if liquidation costs $(\alpha_1 \Delta)$ are sufficiently high, and/or the costs of bailouts ($f(r_0)$ and $f(2r_0)$) are not too steep, then there are "too many (banks) to fail" and the regulator is forced to rescue some or all of the failed banks.

Lemma 2 *In any sub-game perfect equilibrium of state FF, the regulator takes the following actions:*

(i) *If $\alpha_1 \Delta < f(r_0)$, then both banks' assets are sold to outsiders.*

(ii) *If $f(r_0) < \alpha_1 \Delta < f(2r_0) - f(r_0)$, then the regulator bails out one of the banks that is chosen randomly with equal probability. The other bank's assets are sold to outsiders.*

(iii) *If $\alpha_1 \Delta > f(2r_0) - f(r_0)$, then the regulator bails out both banks.*

Furthermore, when a bank is bailed out, the regulator takes a share in the bank's equity of $\beta \le (1 - \bar{\theta})$, but is indifferent between shares over the range $[0, (1 - \bar{\theta}))$.

Thus, the expected second-period profits of the bank depend on the regulator's decision as:

$$E(\pi_2^{ff}) = \begin{cases} 0 & if & \alpha_1 \Delta < f(r_0) \\ \frac{1}{2}(1 - \beta)(\alpha_1 R - 1) & if & f(r_0) < \alpha_1 \Delta < f(2r_0) - f(r_0) \\ (1 - \beta)(\alpha_1 R - 1) & if & \alpha_1 \Delta > f(2r_0) - f(r_0) \end{cases}. \tag{11}$$

3.4 First investment problem (date 0) and inter-bank correlation

In the first period, both banks are identical. Hence, we consider a representative bank. Formally, the objective of each bank is to choose the level of inter-bank correlation ρ at date 0 that maximizes

$$E(\pi_1(\rho)) + E(\pi_2(\rho)), \tag{12}$$

where discounting has been ignored since it does not affect any of the results. Recall that if banks invest in different industries, then inter-bank correlation ρ equals 0, else it equals 1. Since the bank pays the depositors the promised return r_0 only if the return on bank loans is high, the expected payoff to the bank at date 0 from its first-period investment, $E(\pi_1)$, is

$$E(\pi_1) = \alpha_0(R - r_0). \tag{13}$$

Note that banks choose the correlation after deposits are borrowed. Hence, $E(\pi_1)$ does not depend on the level of inter-bank correlation, and banks only take into account the second-period profits when choosing ρ.

We can calculate the expected second-period return of Bank A (and by symmetry, of Bank B) as

$$E(\pi_2(\rho)) = \sum_i \Pr(i)\ E(\pi_2^i(\rho)) \tag{14}$$

where i represents the possible states, that is, $i \in \{SS, SF, FF\}$.

Note that when banks invest in the same industry, $\Pr(SF) = 0$, so that

$$E(\pi_2(1)) = \alpha_0\ E(\pi_2^{ss}) + (1 - \alpha_0)\ E(\pi_2^{ff}). \tag{15}$$

When banks invest in different industries, from Table 3 we obtain that

$$E(\pi_2(0)) = \alpha_0^2\ E(\pi_2^{ss}) + \alpha_0(1 - \alpha_0)\ E(\pi_2^{sf}(0)) + (1 - \alpha_0)^2\ E(\pi_2^{ff}). \tag{16}$$

From Lemma 1, we obtain that $E(\pi_2^{sf}(0)) = E(\pi_2^{ss}) + \alpha_1\Delta$. Thus, we can write

$$E(\pi_2(0)) = \alpha_0\ E(\pi_2^{ss}) + \alpha_0(1 - \alpha_0)\ (\alpha_1\Delta) + (1 - \alpha_0)^2\ E(\pi_2^{ff}) \tag{17}$$

which gives us

$$E(\pi_2(1)) - E(\pi_2(0)) = \alpha_0(1 - \alpha_0)\ [E(\pi_2^{ff}) - \alpha_1\Delta]. \tag{18}$$

Thus, the only terms that affect the choice of inter-bank correlation are the discount $(\alpha_1\Delta)$ the surviving bank gets in state SF from buying the failed bank's assets and the subsidy it receives $(E(\pi_2^{ff}))$ from a bailout in state FF. Using equation (11) for $E(\pi_2^{ff})$ as a function of dilution β employed by the regulator, we obtain the following characterization of the best response of banks in choosing inter-bank correlation.

Lemma 3 *Let $\beta_1^* = 1 - \frac{2\alpha_1 \Delta}{(\alpha_1 R - 1)}$ and $\beta_2^* = 1 - \frac{\alpha_1 \Delta}{(\alpha_1 R - 1)}$.*

(i) *If the regulator liquidates both banks in state FF, that is $\alpha_1 \Delta < f(r_0)$, then banks choose the lowest level of correlation, $\rho = 0$.*

(ii) *If the regulator bails out only one bank in state FF, that is $f(r_0) < \alpha_1 \Delta < f(2r_0) - f(r_0)$, then for a given bailout strategy of $\beta < \beta_1^*$, banks choose the highest level of correlation, $\rho = 1$ and for $\beta > \beta_1^*$, they choose the lowest level of correlation, $\rho = 0$.*

(iii) *If the regulator bails out both banks in state FF, that is $\alpha_1 \Delta > f(2r_0) - f(r_0)$, then for a given bailout strategy of $\beta < \beta_2^*$, banks choose the highest level of correlation, $\rho = 1$ and for $\beta > \beta_2^*$, they choose the lowest level of correlation, $\rho = 0$.*

Proof: *Part (i) is trivial, since in this case $E(\pi_2^{ff}) = 0 < \alpha_1 \Delta$. We prove (iii) first. In this case, both banks are bailed out and $E(\pi_2^{ff}) = (1 - \beta)(\alpha_1 R - 1)$. Banks will choose the highest level of correlation if and only if $E(\pi_2^{ff}) > \alpha_1 \Delta$. This condition holds when $\alpha_1 \Delta < (1 - \beta)(\alpha_1 R - 1)$. We can also write this as $\beta < \beta_2^* = 1 - \frac{\alpha_1 \Delta}{(\alpha_1 R - 1)}$.*

We can prove (ii) along the same lines. In this case, only one bank is bailed out and $E(\pi_2^{ff}) = \frac{1}{2}(1 - \beta)(\alpha_1 R - 1)$. Banks will choose the highest level of correlation if and only if $E(\pi_2^{ff}) > \alpha_1 \Delta$. This condition holds when $\alpha_1 \Delta < \frac{1}{2}(1 - \beta)(\alpha_1 R - 1)$, which gives us $\beta < \beta_1^* = 1 - \frac{2\alpha_1 \Delta}{(\alpha_1 R - 1)}.\diamond$

Note that $\beta_1^* < \beta_2^*$. This means that when both banks are bailed out, the range of the regulator's share over which banks choose high correlation is larger compared to the case when only one of the banks is bailed out. This is intuitive since the subsidy banks receive in state FF is larger.

Finally, in equilibrium the deposit rate r_0 is determined based on the payoff for depositors when the bank fails. If the bank succeeds, depositors get paid the promised payment of r_0. If the bank fails, then the effective payment to depositors is always $[\alpha_1(R - \Delta) - 1]$: In state SF (assuming we are considering Bank A), depositors receive $[\alpha_1(R - \Delta) - 1]$ from the sale of bank's assets to outsiders; in state FF, either the bank is closed in which case they again receive this amount from outsiders, or the bank is bailed out in which case the subsidy to depositors net of taxes is $(r_0 - \tau)$ which is assumed to be equal to $[\alpha_1(R - \Delta) - 1]$.

Hence, individual rationality of depositors implies that

$$\alpha_0 r_0 + (1 - \alpha_0)[\alpha_1(R - \Delta) - 1] = 1. \tag{19}$$

Simplifying, we obtain that

$$r_0 = \frac{1}{\alpha_0}[2 - \alpha_1(R - \Delta)] + [\alpha_1(R - \Delta) - 1]. \tag{20}$$

16

Under our maintained assumptions, $[\alpha_1(R - \Delta) - 1] < 1$ (assumption (iv) in Section 3.2). Hence, $[2 - \alpha_1(R - \Delta)] > 0$, and we obtain that the borrowing rate r_0 is decreasing in α_0, the likelihood of the bank's success in the first period.

We combine the results in Lemma 2 and Lemma 3 with the expression for r_0 to characterize the unique subgame perfect equilibrium in terms of the primitive variables of the model, α_0 and α_1, the likelihood of individual success for each bank in the first period and the second period, respectively. The formal statement is contained in the proposition below and it is captured graphically in Figure 2.

Proposition 4 *In the unique subgame perfect equilibrium, regulator does not intervene in states SS, SF and FS, the surviving bank always buys the assets of the failed bank in states SF and FS, and in state FF, we obtain the following: there exist cutoff levels α_0^* and α_0^{**}, such that $\alpha_0^* \le \alpha_0^{**}$ and*

(i) *If $\alpha_0 < \alpha_0^*$, then $\alpha_1\Delta < f(r_0)$, and the regulator liquidates both banks to outsiders. Under this case, banks invest in different industries at date 0.*

(ii) *If $\alpha_0^* < \alpha_0 < \alpha_0^{**}$, then $f(r_0) < \alpha_1\Delta < f(2r_0) - f(r_0)$, and the regulator liquidates one bank's assets to outsiders and bails out the other. The bailed out bank is randomly chosen with equal probability. If the regulator takes a share of β in the bailed out bank where β is less than $\beta_1^* = 1 - \frac{2\alpha_1\Delta}{(\alpha_1 R - 1)}$, then banks invest in the same industry at date 0; otherwise, they invest in different industries.*

(iii) *If $\alpha_0 > \alpha_0^{**}$, then $\alpha_1\Delta > f(2r_0) - f(r_0)$, and the regulator bails out both banks. If the regulator takes a share of β in the bailed out banks that is less than $\beta_2^* = 1 - \frac{\alpha_1\Delta}{(\alpha_1 R - 1)}$, then banks invest in the same industry at date 0; otherwise, they invest in different industries.*

Furthermore, (i) if $f(2) - f(1) < \alpha_1\Delta$, then $0 < \alpha_0^* < \alpha_0^{**} < 1$, $\alpha_1\Delta = f(r_0(\alpha_0^*))$, $\alpha_1\Delta = f(2r_0(\alpha_0^{**})) - f(r_0(\alpha_0^{**}))$; (ii) if $f(1) < \alpha_1\Delta < f(2) - f(1)$, then $0 < \alpha_0^* < \alpha_0^{**} = 1$, $\alpha_1\Delta = f(r_0(\alpha_0^*))$,; and, finally, (iii) if $f(1) > \alpha_1\Delta$, then $\alpha_0^* = \alpha_0^{**} = 1$.

Note that there is an indeterminacy in our model in the ex-post choice of β, the share that the regulator takes in a bailed-out bank. We analyze below the expectation at date 0 of the output of the banking sector and illustrate that this indeterminacy is resolved when viewed from an ex-ante standpoint for the regulator.

3.5 Time-inconsistency of ex-ante optimal regulation

We show below that the total expected output at date 0 depends on whether banks invest in the same industry or in different industries. However, the regulator cannot write contracts that "force" banks to adopt specific investment choices, that is, the regulator cannot impose regulation that is explicitly contingent on inter-bank correlation.

Let $E(\Pi_t(\rho))$ be the expected output generated by the banking sector in period starting at date t, net of liquidation and/or bailout costs. If banks invest in the same industry at date 0, then with probability α_0 both banks have the high return so that $E(\Pi_1(1)) = 2\alpha_0 R$. However, if they invest in different industries, then with probability α_0^2 both banks have the high return whereas with probability $2\alpha_0(1 - \alpha_0)$, one bank has the high return while the other has the low return. This gives us $E(\Pi_1(0)) = 2\alpha_0^2 R + 2\alpha_0(1 - \alpha_0)R = 2\alpha_0 R$. Thus, total expected output in the first period is independent of the choice of inter-bank correlation.

In the second period, the number of banks that operate depends on the outcome of the first-period investments and the regulator's action. In state SS, both banks operate one more period. In state SF or FS, failed bank's assets are purchased by the surviving bank and the total expected output is the same as the total expected output in state SS, that is, $E(\Pi_2^{ss}(\rho)) = E(\Pi_2^{sf}(\rho)) = E(\Pi_2^{fs}(\rho)) = 2\alpha_1 R$, for $\rho \in \{0, 1\}$.

However, the expected output in state FF depends on the closure policy adopted by the regulator. In this state, either banks are closed and liquidation costs are incurred, or, some banks are bailed out by the regulator at a fiscal cost. In either case, there is a social welfare loss relative to the other states. Thus, $E(\Pi_2^{ff}(\rho)) < E(\Pi_2^{ss}(\rho))$, for $\rho \in \{0, 1\}$. Using the corresponding joint probabilities, we get:

$$E(\Pi_2(1)) = \alpha_1 E(\Pi_2^{ss}) + (1 - \alpha_1)E(\Pi_2^{ff}), \tag{21}$$
$$E(\Pi_2(0)) = \alpha_1^2 E(\Pi_2^{ss}) + 2\alpha_1(1 - \alpha_1)E(\Pi_2^{sf}) + (1 - \alpha_1)^2 E(\Pi_2^{ff}). \tag{22}$$

Note that $\alpha_1^2 E(\Pi_2^{ss}) + \alpha_1(1 - \alpha_1)E(\Pi_2^{sf}) = \alpha_1 E(\Pi_2^{ss})$, so that,

$$E(\Pi_2(0)) - E(\Pi_2(1)) = \alpha_1(1 - \alpha_1)\left[E(\Pi_2^{ss}) - E(\Pi_2^{ff})\right] > 0. \tag{23}$$

This gives us the following result:

Lemma 5 *Expected total output of the banking sector at date* 0 *(net of any anticipated costs of liquidations and bailouts) is maximized when banks operate in different industries, that is, when $\rho = 0$.*

Since FF is the only state where there is a social welfare loss, the regulator may wish to implement closure policies that minimize the probability of state FF, that is, policies

that give incentives for banks to choose low correlation. These policies may however not be ex-post optimal. For example, committing to liquidate both banks in state FF has the ex-ante advantage that it gives banks incentives to invest in different industries. However, conditional upon reaching state FF, liquidation of banks may not be credible if costs of bail out are smaller than liquidation costs. Another way the regulator can induce low correlation among banks is by diluting the equity share of bailed-out banks in state FF (see Lemma 3). However, this may also lack commitment ex post: if the minimum dilution required to induce low correlation is sufficiently large, then such dilution may have adverse consequences for continuation moral hazard and banks may choose bad projects.

We formalize this trade-off below. In particular, we characterize the ex-ante optimal regulatory policy assuming that the regulator can commit to ex-post implementation of this policy. We also examine the case where the ex-ante optimal policy is not subgame perfect and thus time-inconsistent.

Consider the three cases for state FF as in Lemma 3 and Proposition 4.

In the first case, we have $\alpha_1 \Delta < f(r_0)$ (or in other words, $\alpha_0 < \alpha_0^*$), and it is ex-post optimal to liquidate both banks. That is, $E(\Pi_2^{ff})$ is maximized by liquidating banks to outsiders. In turn, this also induces banks to invest in different industries. Hence, it is also ex-ante optimal to commit to liquidating both banks in state FF.

In the more interesting second case, we have $f(r_0) < \alpha_1 \Delta < f(2r_0) - f(r_0)$ (or in other words, $\alpha_0^* < \alpha_0 < \alpha_0^{**}$) and it is ex-post optimal to liquidate one bank and bail out the other bank. Ex ante, the regulator wishes to implement this ex-post optimal outcome and yet induce a low correlation among banks at date 0. The regulator can achieve this if it can take a share $\beta > \beta_1^*$ in the bailed-out bank without inducing continuation moral hazard. That is, β should be greater than β_1^* (as defined in Lemma 3) to induce low correlation, but be smaller than $(1 - \overline{\theta})$ in order to provide continuation incentives. If $\beta_1^* < (1 - \overline{\theta})$, then such a dilution scheme can be implemented by choosing $\beta = \beta_1^*$ and it is ex-post credible.

However, if $\beta_1^* > (1 - \overline{\theta})$, then a dilution scheme that sets $\beta = \beta_1^*$ is dominated ex ante by a strategy that liquidates the bank instead. This is because under our maintained assumption $(\overline{\Delta} > \Delta)$, liquidation costs of a bank are smaller than agency costs arising from an excessive dilution of bank owners' stake in profits. Is it ex-ante optimal for the regulator to commit to liquidate *both* banks in this case even though it is ex-post optimal to liquidate *only one* of the banks and bail out the other? The answer is yes for at least a part of the parameter range. To see this, note that if the regulator can commit ex ante to liquidating both banks, this induces banks to choose the low correlation. Therefore, in this case $E(\Pi_2^{ff}) = E(\Pi_2^{ss}) - 2\alpha_1 \Delta$. In turn, this gives

$$E(\Pi_2(0)) = E(\Pi_2^{ss}) - (1 - \alpha_0)^2 (2\alpha_1 \Delta), \tag{24}$$

since state FF is likely with a probability of $(1-\alpha_0)^2$ when banks invest in different industries.

If the regulator commits to the ex-post optimal strategy of liquidating only one of the banks, then it cannot implement a low correlation. Hence, $E(\Pi_2^{ff}) = E(\Pi_2^{ss}) - \alpha_1\Delta - f(r_0)$, and

$$E(\Pi_2(1)) = E(\Pi_2^{ss}) - (1 - \alpha_0)\left[\alpha_1\Delta + f(r_0)\right], \tag{25}$$

since state FF is likely with a probability of $(1-\alpha_0)$ when banks invest in the same industry.

Thus, it is optimal for the regulator to commit to liquidating both banks if and only if $E(\Pi_2(0)) > E(\Pi_2(1))$, a condition that can be simplified to, $(1 - 2\alpha_0)\alpha_1\Delta < f(r_0)$. From Proposition 4, this second case where $f(r_0) < \alpha_1\Delta < f(2r_0) - f(r_0)$ holds, arises if α_0 is between the cutoff levels α_0^* and α_0^{**}. Note that at α_0^*, $\alpha_1\Delta = f(r_0(\alpha_0))$, and in turn, $[(1 - 2\alpha_0)\alpha_1\Delta - f(r_0)] = -2\alpha_0\alpha_1\Delta < 0$. Hence, it follows that $(1 - 2\alpha_0)\alpha_1\Delta < f(r_0)$ for at least a part of the range $[\alpha_0^*, \alpha_0^{**}]$ including α_0^*. We denote this range over which the regulator would like to commit to liquidating both banks as $[\alpha_0^*, \hat{\alpha}_0]$.

In other words, the regulator would like to commit ex-ante to liquidating both banks over a larger range of α_0 compared to the range over which it is ex-post optimal to liquidate both banks. The trade-off is simple: ex post, the regulator cares only about expected profits in state FF, whereas ex ante, the regulator is willing to give up some of these profits in order to induce better incentives for banks to be less correlated and reduce the likelihood of ending up in state FF. We can use a similar analysis (see Appendix for the proof) for the case where $\alpha_1\Delta < f(2r_0) - f(r_0)$ holds and it is ex-post optimal for the regulator to bail out both banks. This gives us the following proposition on the time-inconsistency of ex-ante optimal regulation. The range of the primitive parameter α_0 over which time-inconsistency arises is illustrated graphically in Figure 3.

Proposition 6 *The ex-ante optimal regulation that maximizes the expected output of the banking sector at date 0 is as follows: in states SS, SF, and FS, the regulator does not intervene and this is also ex-post optimal; in state FF, the regulator's ex-ante optimal policy is time-inconsistent when the bank-level moral hazard is sufficiently severe:*

(1) *$(1 - \bar{\theta}) < \beta_1^*$: There exists a cutoff $\hat{\alpha}_0$ such that $\alpha_0^* < \hat{\alpha}_0$, where α_0^* is as characterized in Proposition 4, and for $\alpha_0^* < \alpha_0 < \hat{\alpha}_0$, the regulator commits ex ante to liquidating both banks to outsiders and banks invest in different industries at date 0. Ex post, it is optimal to bail out at least one bank and this induces banks to invest in the same industry at date 0.*

(2) *$\beta_1^* < (1 - \bar{\theta}) < \beta_2^*$: There exists a cutoff $\hat{\hat{\alpha}}_0$ such that $\alpha_0^{**} < \hat{\hat{\alpha}}_0 \leq 1$, where α_0^{**} is as characterized in Proposition 4, and for $\alpha_0^{**} < \alpha_0 < \hat{\hat{\alpha}}_0$, the regulator commits ex ante*

to liquidating one (randomly chosen) bank to outsiders and taking a share β_1^* in the bailed-out bank. This induces banks to invest in different industries at date 0. Ex post, it is optimal to bail out both banks and this induces banks to invest in the same industry at date 0.

In other cases, it is possible for the regulator to implement a low correlation among banks at date 0 without affecting the ex-post optimal closure policy. This is because when the bank-level moral hazard is small, the regulator is able to dilute bailed-out bank(s)' equity sufficiently to make the bailout subsidy in joint-failure state smaller than the surplus gained by banks in individual survival state. This induces banks to invest in different industries, and crucially, without exacerbating incentives in the continuation game.

3.6   Bank sales at discount prices and partial nationalization

It is useful to consider the efficacy of two tools that the regulator may be able to employ in order to counteract the moral hazard of bailout subsidies. In the model, the decision by banks to herd is determined by the relative levels of (i) the surplus captured in the individual survival state (SF) by acquiring the failed bank at a discount of $\alpha_1\Delta$, and (ii) the bailout subsidy in the joint-failure state (FF) that depends upon the extent of nationalization of the bailed-out banks, represented by the share β taken by the regulator. This suggests that perhaps the regulator could intervene in state SF as well and allow the surviving bank to purchase the failed back at a further discount.

To see this, suppose that the failed bank is sold to the surviving bank at a price $[\alpha_1(R - \Delta') - 1]$ where $\Delta' > \Delta$, that is, at an opportunity cost to creditors of the failed bank who can sell the bank to outsiders at a better price of $[\alpha_1(R - \Delta) - 1]$. In this case, the surviving bank purchases the failed bank at a discount of $\alpha_1\Delta'$ (greater than the discount $\alpha_1\Delta$ in our benchmark model). Then, in Lemma 3, the threshold levels β_1^* and β_2^* become smaller, and the regulator can implement a low correlation among banks by taking smaller shares in bailed-out banks in state FF. In other words, if it is possible to force losses on depositors of failed banks when only a few banks have failed, and partially (but not excessively) nationalize bailed-out banks when a large number of banks have failed, then the regulator can mitigate the systemic risk induced by its bailout policy.

While we believe this is an interesting prescription for bank-closure policy, it also needs to be qualified with some caveats. First, in anticipation of regulatory intervention on the prices at which bank sales can take place, failed banks may advance their sales to outside investors. Regulatory intervention in bank sales when banks have not yet defaulted may be infeasible as this would be viewed as too intrusive by the banking sector. Second, even

if such intervention were feasible, forcing losses on depositors would raise the equilibrium cost of borrowing deposits, and make it more costly and difficult to bail out banks ex post (if it is so required in state FF). Though this may alleviate incentives to herd ex ante, it increases ex-post costs of banking crises. Finally, the extent of moral hazard (measured by $\bar{\theta}$ in the model) is unobservable and it may be impossible for the regulator to know (and commit to enforcing) the optimal extent of nationalization in joint-failure states. The precise choices in an implementation of these mechanisms may thus be a tricky calibration exercise for prudential regulators.[6]

3.7 Robustness issues

It is important to note that a number of assumptions we made to derive our results are biased against generating a too-many-to-fail effect and/or the induced herding behavior of banks. First, we assumed that in state SF the surviving bank has all the bargaining power and can make a take-it-or-leave-it offer at $[\alpha_1(R-\Delta)-1]$ to the failed bank. If the surplus of $\alpha_1\Delta$ was instead split between the two banks, then the strategic benefit from surviving individually would be even smaller. In turn, the cutoff levels for share taken by the regulator below which banks herd, β_1^* and β_2^*, would be even larger.

Second, we assumed that depositors of bailed-out banks are taxed in state FF such that depositors did not receive a net subsidy in state FF relative to state SF, and the cost of borrowing at date 0, r_0, did not depend on inter-bank correlation. Suppose there was a net subsidy for depositors in state FF relative to state SF, then r_0 would be smaller when inter-bank correlation is high. This in turn would reduce the funds needed for bailouts in state FF, make bailouts more likely, and reinforce herding by banks. Finally, we assumed that the cost of providing funds with immediacy for bailing out banks, $f(x)$, is convex in the amount of funds x. If we assumed the cost function was linear, then depending on the magnitude of liquidation costs, $\alpha_1\Delta$, either both banks would be liquidated or both banks would be bailed out. Under the assumption of convex costs, it is sometimes optimal to bail out just one bank and not both, attenuating the too-many-to-fail problem and the induced herding.

There are also considerations outside the model that lend an element of robustness to the too-many-to-fail problem and the welfare costs arising from the induced herding behavior.

[6]An alternative choice available to the regulator is to allow a private-sector resolution of the failed bank at the "market" price of $[\alpha_1(R-\Delta)-1]$, but to transfer a subsidy to the surviving bank in the form of direct injection of capital. Such "relative-performance reward" would entail fiscal costs to the regulator, but can provide incentives to banks not to herd. While this mechanism appears feasible, we believe that it is not robust to other considerations. For instance, institutionalizing regulatory capital injections into surviving banks can lead to or exacerbate other time-inconsistency issues related to lobbying by surviving banks to get funds, favoring of state-owned banks (if any), etc.

For instance, bank closure policy is often marred by considerations of regulatory reputation (as witnessed during the S&L crisis in the United States, and as modeled by Boot and Thakor, 1993) and political economy (as shown empirically by Brown and Dinc, 2004, from a study linking delay in bank failure announcements to elections in 21 major emerging markets). In the presence of such effects, the short horizon of regulatory decision-makers may lead them to exercise forbearance and bail out banks when it is not even ex-post optimal. This is clearly more likely if many banks have failed. Considerations of bank competition may also make it difficult for the regulator to allow failed banks to be acquired by surviving banks when there are only a few banks left. This would render it difficult for the regulator to counteract the perverse effect of bailout subsidies on ex-ante herding incentives.

Similarly, the welfare losses from bank liquidations in joint-failure states may arise not just from an allocation inefficiency as in our model, but possibly also from the loss of consumer confidence, contagious runs on other banks (see Allen and Gale, 2000, and references therein), disruptions in credit creation and investments, problems relating to the payment systems, and accentuation of liquidity problems in the banking system (see Diamond and Rajan, 2003, and the references therein). Some of these costs arise due to banking crises per se, and not because of specific regulatory actions undertaken in these times. In a similar vein, herding may not only increase the likelihood of joint-failure states, but also lead to a bypassing of valuable projects by banks. To this extent, the moral hazard induced by bailouts in joint-failure states may have quite adverse welfare consequences from an ex-ante standpoint.

It is in order to point out that the possibility of banking failures will affect investment decisions of banks significantly only when such failures are sufficiently likely. When such failures are not too likely, countervailing effects such as competition in loan margins may provide sufficient "anti-herding" incentives to banks. Finally, we view the too-many-to-fail channel of herding as being complementary to the herding channels discussed in the literature. These other channels include managerial herding in disclosure of losses based on reputation considerations (Rajan, 1994) and herding by banks to avail cheaper borrowing rates when they perform well together (Acharya and Yorulmazer, 2003). Overall, we view our channel of herding as being more specific to the banking sector than these channels since it is based on regulatory subsidies.

4   Conclusion

This paper makes a simple but a fundamental point that there is an inherent time-inconsistency in bank closure policies. In a systemic crisis, that is when many banks fail, the regulator has no ex-post choice but to bail out some or all of these banks. However, this is inefficient ex ante since it gives banks incentives to herd and increases the likelihood of systemic crises.

We call this a "too-many-to-fail" problem. The literature and central banks have primarily focused on the "too-big-to-fail" problem which has been explicitly recognized by bank regulators.[7] In contrast, the too-many-to-fail guarantees have not been explicitly recognized even though they have been provided regularly to banks during systemic banking crises. Furthermore, these guarantees accrue not only to the very large banks (as in the too-big-to-fail guarantee) but also to other, relatively smaller banks. Hence, we believe that our analysis of the (sub-)optimality of these guarantees and their ex-ante costs is important and novel from an academic standpoint as well as from a policy standpoint. The main implication of our analysis is that the genesis of inefficient systemic risk may potentially lie in the very crises-management role of central banks or equivalent bank regulators. Thus, the paper highlights the need for understanding and designing regulatory policies (i) at a systemic level rather than only at an individual bank level (as is the current practice), and (ii) at a crises-prevention level rather than only at a crises-management level.

---

[7]See, for example, O'Hara and Shaw (1990) who study the effect of the public announcement of too-big-to-fail guarantees by the Office of the Comptroller of the Currency in the United States.

References

Acharya, Viral (2001) A Theory of Systemic Risk and Design of Prudential Bank Regulation, Working Paper, New York University and London Business School.

Acharya, Viral and Tanju Yorulmazer (2003) A Theory of Procyclical Bank Herding, Working Paper, London Business School.

Allen, Franklin and Douglas Gale (2000) Financial Contagion, *Journal of Political Economy*, 108, 1–33.

Bagehot, William (1873) *Lombard Street: A Description of the Money Market*, London: H. S. King.

Boot, Arnoud and Anjan Thakor (1993) Self-interested Bank Regulation, *American Economic Review*, 83(2), 206–212.

Brown, Craig O'Neil and I. Serdar Dinc (2004), The Politics of Bank Failures: Evidence from Emerging Markets, *Journal of Financial Economics*, forthcoming.

Diamond, Douglas and Raghuram Rajan (2003) Liquidity Shortages and Banking Crises, *Journal of Finance*, forthcoming.

Freixas, Xavier (1999) Optimal Bail out Policy, Conditionality and Constructive Ambiguity, Working Paper, Universitat Pompeu Fabra.

Gorton, Gary (1985) Clearinghouses and the Origin of Central Banking in the United States, *Journal of Economic History*, XLV(2), 277–283.

Hart, Oliver and John Moore (1994) A Theory of Debt Based on the Inalienability of Human Capital, *Quarterly Journal of Economics*, 109(4), 841–879.

Hoggarth, Glenn, Reis, Ricardo and Victoria Saporta (2001) Costs of Banking System Instability: Some Empirical Evidence, Working Paper, Bank of England.

Hoggarth, Glenn, Reidhill, Jack and Peter Sinclair (2004) On the Resolution of Banking Crises: Theory and Evidence, Working Paper #229, Bank of England.

Holmstrom, Bengt and Jean Tirole (1998) Private and Public Supply of Liquidity, *Journal of Political Economy*, 106(1), 1–40.

James, Christopher (1991) The Losses Realized in Bank Failures, *Journal of Finance*, 46(4), 1223–

Kasa, Kenneth and Mark Spiegel (1999) The Role of Relative Performance in Bank Closure Decisions, Working Paper, Federal Reserve Bank of San Francisco.

Mailath, George and Loretta Mester (1994) A Positive Analysis of Bank Closure, *Journal of Financial Intermediation*, 3, 272–299.

Mitchell, Janet (1997) Strategic Creditor Passivity, Bank Rescues, and Too Many to Fail, CEPR Working Paper #1780.

Myers, Stewart and Nicholas Majluf (1984) Corporate Financing and Investment When Firms Have Information Shareholders Do Not Have, *Journal of Financial Economics*, 13, 187–221.

O'Hara, Maureen and Wayne Shaw (1990) Deposit Insurance and Wealth Effects: The Value of Being "Too Big to Fail," *Journal of Finance*, 45(6), 1587–1600.

Penati, Alessandro and Aris Protopapadakis (1988) The Effect of Implicit Deposit Insurance on Banks' Portfolio Choices, with an Application to International 'Overexposure', *Journal of Monetary Economics*, 21(1), 107—126.

Perotti, Enrico and Javier Suarez (2002) Last Bank Standing: What do I Gain If You Fail?, *European Economic Review*, 46(09), 1599–1622.

Rajan, Raghuram (1994) Why Bank Credit Policies Fluctuate: A Theory and Some Evidence, *Quarterly Journal of Economics*, 109, 399–442.

Santomero, Anthony and Paul Hoffman (1999) Problem Bank Resolution: Evaluating the Options, Working Paper, Financial Institutions Center, the Wharton School, University of Pennsylvania.

Appendix

Proof of the second case in Lemma 1: This is the case in state SF where depositors can get their promised return only when both banks' assets perform well. The distribution of bank profits in this case is given in equation (5). For this case to hold, we need $R < 2r_1^{sf}$ where r_1^{sf} is the deposit rate the bank needs to promise to attract depositors. We assume that when the bank cannot pay depositors their promised return, depositors equally share whatever is available in the bank and the bank owners get nothing. Therefore, depositors get the promised r_1^{sf} when the return is (R, R) and they get nothing when the return is $(0, 0)$. In states $(R, 0)$ and $(0, R)$, the return from the pooled assets is R and that is distributed equally among depositors. Thus, each set of depositors receive $R/2$.

Using these payoffs and the individual rationality of depositors, we can calculate the promised deposit rate as

$$\Pr(R, R) \ r_1^{sf} + 2\Pr(R, 0) \ (R/2) = 1. \tag{26}$$

This gives us $\alpha_1^2 \ r_1^{sf} + \alpha_1(1 - \alpha_1)R = 1$. So we get $r_1^{sf} = \frac{[1 - \alpha_1(1 - \alpha_1)R]}{\alpha_1^2}$.

Next, we need to show that the bank prefers to purchase the other bank's assets. With the asset purchase, the expected profit from the second-period investment is:

$$E(\pi_2^{sf}(0)) \ = \ \Pr(R, R) \ (2R - 2r_1^{sf}) - [\alpha_1(R - \Delta) - 1] \tag{27}$$
$$= \ \alpha_1^2(2R - 2r_1^{sf}) - [\alpha_1(R - \Delta) - 1]. \tag{28}$$

Plugging in r_1^{sf} in this equation and simplifying we get:

$$E(\pi_2^{sf}(0)) \ = \ (\alpha_1 R - 1) + \alpha_1\Delta. \tag{29}$$

Without the asset purchase, expected profit from second investment is $(\alpha_1 R - 1)$. Therefore, the surviving bank always purchases the failed banks assets and makes a profit of $\alpha_1\Delta$. $\diamondsuit$

Proof of Proposition 4: Note that the proposition follows from Lemma 2 and Lemma 3, except that the three conditions have to be derived in terms of α_0.

Denote $g(\alpha_0) = f(r_0(\alpha_0))$. From equation (20) for r_0, we obtain that $g'(\alpha_0) = f'(r_0)\frac{\partial r_0}{\partial \alpha_0} < 0$ since $f'(r_0) > 0$ and r_0 is decreasing in α_0. Furthermore, $g(0) \to \infty$. Thus, if $g(1) = f(r_0(1)) = f(1)$ is smaller than $\alpha_1\Delta$, then there exists a unique α_0^* such that $0 < \alpha_0^* < 1$ and $g(\alpha_0^*) = \alpha_1\Delta$. Otherwise, $\alpha_0^* = 1$. It follows then that $\alpha_1\Delta < f(r_0)$ if and only if $\alpha_0 < \alpha_0^*$.

Next, note that since $f(r_0)$ is a convex function, we have that $f(2r_0) > 2f(r_0)$ or in other words $f(2r_0) - f(r_0) > f(r_0)$. Denote $h(\alpha_0) = f(2r_0(\alpha_0)) - f(r_0(\alpha_0))$. Then, $h(\alpha_0) > g(\alpha_0)$ for all α_0. Furthermore, $h(0) \to \infty$ and $h'(\alpha_0) = [2f'(2r_0) - f'(r_0)]\frac{\partial r_0}{\partial \alpha_0} < 0$. Thus, if

$h(1) = f(2) - f(1)$ is smaller than $\alpha_1 \Delta$, then there exists a unique α_0^{**} such that $0 < \alpha_0^{**} < 1$ and $g(\alpha_0^{**}) = \alpha_1 \Delta$. Otherwise, $\alpha_0^{**} = 1$. It follows then that $\alpha_1 \Delta < f(2r_0) - f(r_0)$ if and only if $\alpha_0 < \alpha_0^{**}$. Finally, since $h(\alpha_0) > g(\alpha_0)$, we must have $\alpha_0^* \leq \alpha_0^{**}$ (with strict inequality when $\alpha_0^* < 1$). $\diamondsuit$

Proof of Proposition 6: In the third case, we have $\alpha_1 \Delta > f(2r_0) - f(r_0)$ (or in other words, $\alpha_0 > \alpha_0^{**}$) and it is ex-post optimal to bail out both banks. Ex ante, the regulator wishes to implement this ex-post optimal outcome and yet induce a low correlation among banks at date 0. The regulator can achieve this if it can take a share $\beta > \beta_2^*$ (as defined in Lemma 3) to induce low correlation, but still have $\beta < (1 - \overline{\theta})$ in order to provide continuation incentives. If $\beta_2^* < (1 - \overline{\theta})$, then such a dilution scheme can be implemented by choosing $\beta = \beta_2^*$ and it is ex-post credible.
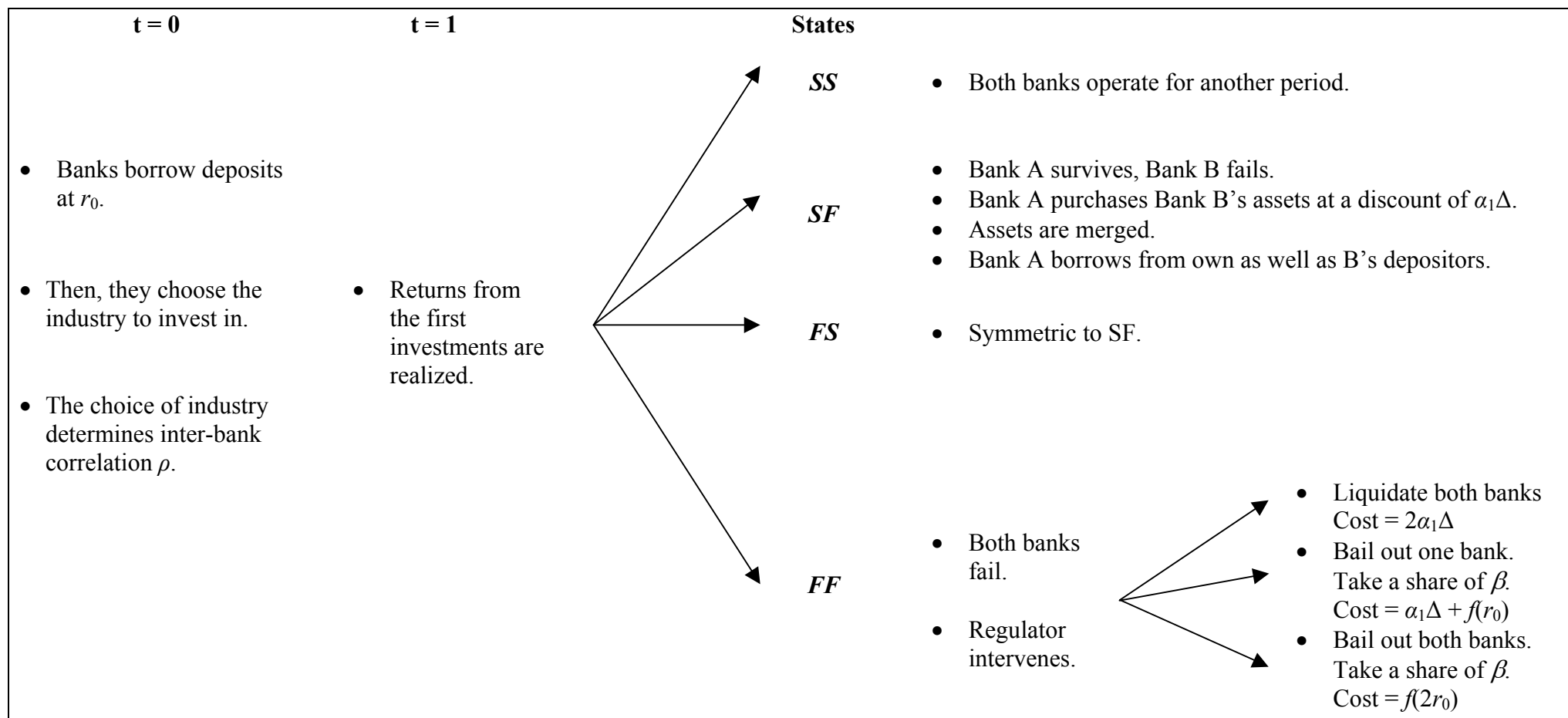
However, if $\beta_2^* > (1 - \overline{\theta})$, then a dilution scheme that sets $\beta = \beta_2^*$ is dominated ex ante by a strategy that liquidates the bailed-out banks instead. Suppose that $\beta_2^* > (1 - \overline{\theta}) > \beta_1^*$. Note that this strategy dominates one of liquidating both banks since liquidating only one bank results in lower ex-post costs but still implement a low correlation. Then, the regulator can implement a low correlation by committing to liquidate only one of the banks and diluting the share of the bailed-out bank by $\beta = \beta_1^*$. We show that it would be ex-ante optimal for the regulator to commit to do this (at least for some parameter range) even though it is ex-post optimal to bail out both banks. Note that if the low correlation is implemented by committing to liquidate one bank, then $E(\Pi_2^{ff}) = E(\Pi_2^{ss}) - \alpha \Delta - f(r_0)$, and

$$E(\Pi_2(0)) = E(\Pi_2^{ss}) - (1 - \alpha_0)^2 \left[\alpha_1 \Delta + f(r_0)\right]. \tag{30}$$
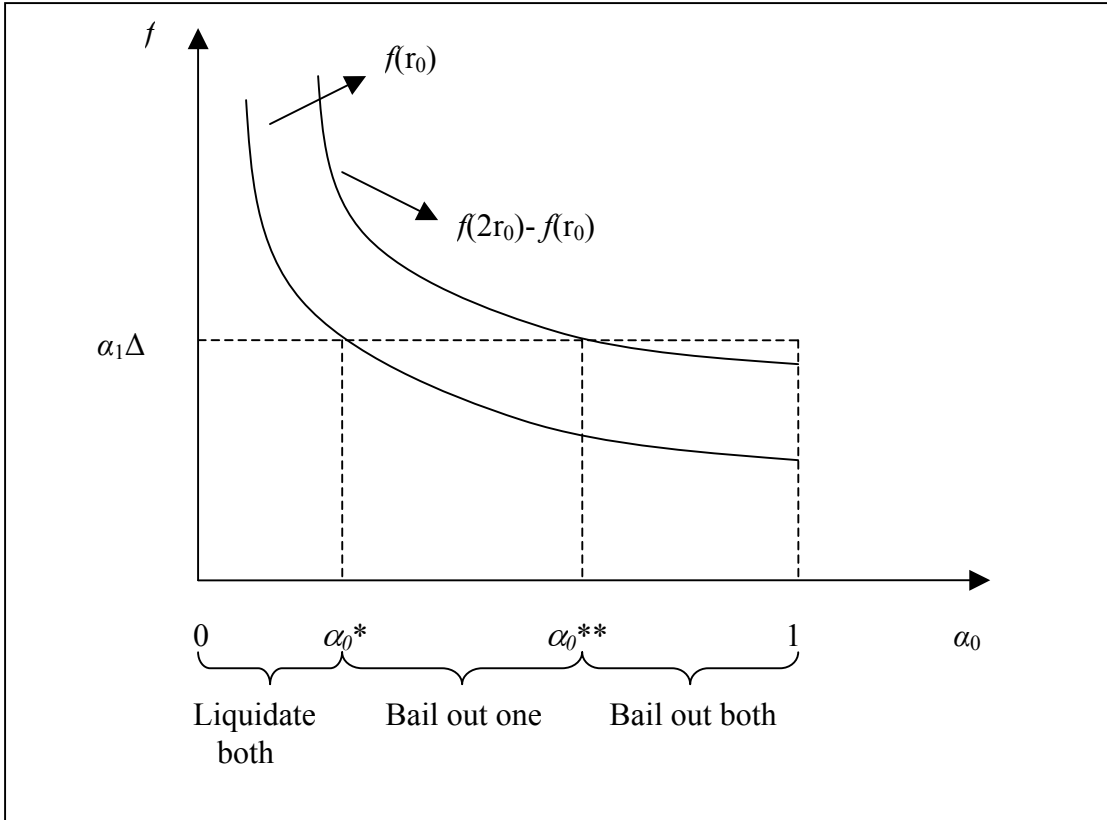
Instead, if the regulator commits to the ex-post optimal strategy of bailing out both banks, then it cannot implement a low correlation. Hence, $E(\Pi_2^{ff}) = E(\Pi_2^{ss}) - f(2r_0)$, and

$$E(\Pi_2(1)) = E(\Pi_2^{ss}) - (1 - \alpha_0)f(2r_0). \tag{31}$$

Thus, it is optimal for the regulator to commit to liquidating one of the banks if and only if $E(\Pi_2(0)) > E(\Pi_2(1))$, a condition that can be simplified to, $(1 - \alpha_0)[\alpha_1 \Delta + f(r_0)] < f(2r_0)$. Note that at α_0^{**}, $\alpha_1 \Delta = f(2r_0) - f(r_0)$, and in turn, $(1 - \alpha_0)[(\alpha_1 \Delta + f(r_0)] - f(2r_0) = -\alpha_0 f(2r_0) < 0$. Hence, it follows that $(1 - \alpha_0)[\alpha_1 \Delta + f(r_0)] < f(2r_0)$ for at least a part of the range to the right of α_0^{**} including α_0^{**}. We denote this range as $[\alpha_0^{**}, \hat{\alpha}_0]$. $\diamondsuit$

t = 0	t = 1	States

SS
- Both banks operate for another period.

- Banks borrow deposits at r_0.

SF
- Bank A survives, Bank B fails.
- Bank A purchases Bank B's assets at a discount of $\alpha_1\Delta$.
- Assets are merged.
- Bank A borrows from own as well as B's depositors.

- Then, they choose the industry to invest in.

- Returns from the first investments are realized.

FS
- Symmetric to SF.

- The choice of industry determines inter-bank correlation ρ.

FF
- Both banks fail.
- Regulator intervenes.

- Liquidate both banks
  Cost = $2\alpha_1\Delta$
- Bail out one bank. Take a share of β.
  Cost = $\alpha_1\Delta + f(r_0)$
- Bail out both banks. Take a share of β.
  Cost = $f(2r_0)$

Figure 1: Timeline of the model.

Figure 2: Ex-post optimal closure policy in state *FF* (Proposition 3.4)
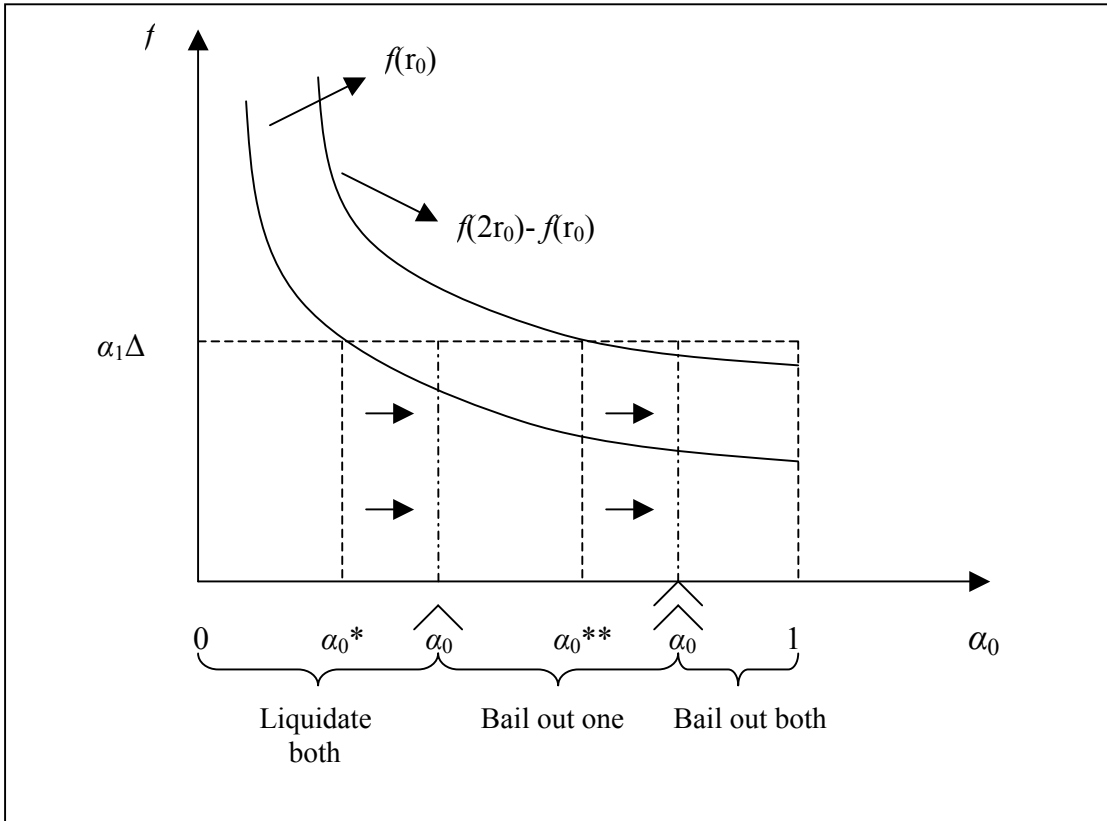
Figure 3: Time-inconsistency of ex-ante closure policy (Proposition 3.6)