Let the Worst One Fail: A Credible Solution to the Too-Big-To-Fail Conundrum*

Thomas Philippon[†] Olivier Wang[‡]

December 2021

Abstract

We study time-consistent bank resolution mechanisms. When interventions are ex post efficient, a government cannot commit not to inject capital into the banking system. Contrary to common wisdom, we show that the government may still avoid moral hazard and implement the first best allocation by using the distribution of bailouts across banks to provide ex ante incentives. We analyze properties of credible tournament mechanisms that provide support to the best performing banks and resolve the worst performing ones, including through mergers. Our mechanism continues to perform well when banks are heterogeneous in size, when they are imperfect substitute, and when they are differentially interconnected as long as bailout funds can be earmarked.

*We are grateful to our discussants Jean-Edouard Colliard, Jason Donaldson, and Jeremy Stein, to Jean Tirole for his many insightful comments, and to participants at the 2021 Adam Smith Workshop, the 2021 NYU-NY Fed Conference, the 2021 NBER Summer Institute, Johns Hopkins, WUSTL, Harvard, Princeton, HEC, Toulouse, and Rutgers.

[†]New York University Stern School of Business, CEPR and NBER, tphilipp@stern.nyu.edu

[‡]New York University Stern School of Business, olivier.wang@nyu.edu

Governments often bail out large financial firms during financial crises because they perceive that the economic costs of letting these firms fail exceed the fiscal costs of the bailouts themselves. This recurrent issue came to a head during the global financial crisis (GFC) of 2008-2009 because of the magnitude and scope of the bailouts. In the aftermath of the Great Recession, governments pledged to end the "too-big-to-fail" problem, and G20 Leaders endorsed the global implementation of a set of reforms for systemically important banks (SIBs). These financial stability reforms rely on three pillars: capital requirements (and other forms of loss absorbing capacity), enhanced supervision, and resolution regimes. The reforms have achieved significant progress along the first two dimensions. Capital requirements have roughly doubled and the supervision of large banks has become tighter (Financial Stability Board, 2021). These evolutions are somewhat uneven across jurisdictions, but regulators and market participants view banks as significantly safer than before the GFC.

The same cannot be said, however, of the third pillar: resolution regimes. Despite 10 years of efforts, there is still no consensus about the ability of governments to resolve large banks during times of economic stress. The root of the skepticism is that one cannot expect policy makers to let a majority of banks – or even a significant number of large ones – fail at the same time. As a result, the argument goes, the expectation of bailouts will remain and will continue to distort funding costs and to feed moral hazard.

We argue that this skepticism is misplaced. More precisely, while we agree with the premise (letting several large banks fail is not a realistic option), we show that the pessimistic conclusion does not follow. The logic of the standard argument is flawed in two ways. Firstly, it assumes that if regulators cannot let a majority of banks fail then no bank can fail at all. Secondly, it assumes that private incentives depend only on the average level of the bailout. We show that both arguments are incorrect.

The main idea of our paper is to apply the logic of tournaments to the issue of too-big-to-fail in the context of imperfect resolution regimes. We assume that it is impossible for governments to credibly commit not to intervene to support the financial sector as a whole during a crisis. However, this does not mean that the government has to support every bank in the same way. Time consistency might pin down the size of the bailout but it does not generally pin down its distribution, and the distribution of bailout funds (or taxes) matters for incentives.

We write a simple model where bailouts can be *ex post* efficient because of a negative externality on the real economy when the financial system is undercapitalized. Bailout anticipations affect the incentives of banks to engage in costly risk mitigation strategies *ex ante*. When we assume, as in the existing literature, that bailout funds are distributed in a symmetric way across banks, we obtain the standard moral hazard results: bailouts inefficiently increase risk taking as in Chari and Kehoe (2016), create strategic complementarities across banks' risk management choices as in Farhi and Tirole (2012), and the situation is worse the deeper the pockets of the government. This line of argument strongly calls for limiting the funds available for bailouts and tying the hands of regulators ex post to the extent possible.

To establish our first main result we use the systemic risk model of Acharya et al. (2016) where the negative externality on the real economy depends on the aggregate capital shortfall in the banking system. In this case the optimal bailout takes the form of a weakly increasing function $\mathcal{M}(K-R)$ where K is the aggregate capital requirement and R the aggregate return. With N banks, time consistency requires that the set of bailout payments satisfies $\sum_{i=1}^{N} m_i = \mathcal{M}(K-R)$ for any value of $R = \sum_{i=1}^{N} r_i$. This places no restrictions of the distribution of $\{m_i\}$ around its mean. In stark contrast to the conventional results, we then show that we can implement the first best equilibrium by conditioning government support on a relative performance mechanism such as a rank-order tournament, in which banks performing above the median get a higher m than banks performing below the median. The scheme is fully time consistent since it takes as given the overall size of the bailout. Punishing the banks that perform poorly while rewarding those who perform well works because, despite knowing that the median bank will be saved, each individual bank strives to make sure it does not end up in the lower half. This race to the top generate first best ex ante incentives for all the banks.

The optimal contract requires punishment of bad banks. When we extend our model by adding limited liability constraints, we find that the common wisdom regarding deep pockets is overturned. We show that the set of implementable policies improves monotonically with fiscal slack. The more slack, the more incentives the government can provide, the less moral hazard, and with enough slack the first best is always implementable despite limited liability. When the limited liability constraint binds, our model offers a macro-prudential justification for increasing TLAC requirements and also for mandating clawback provisions in executive compensation contracts. The reason is that these contracts reduce the tightness of the constraint and therefore increase the range of time consistent outcomes. For the same reason, we show that although the fire sales that occur during systemic crises must be met by larger bailouts, they also make it easier to provide ex ante incentives. Fire sales hurt the outside options of weak banks relative to the transfers proposed by the regulator. Reducing bank leverage improves risk-taking incentives when fire sales discount are deep enough.

Our baseline framework assumes that banks are highly substitutable, in the sense that capital surpluses in one bank can compensate for capital shortfalls in another. We show that this pure systemic risk model can be viewed as the optimal outcome of a process that allows the resolution authority to merge banks at a low cost. If healthy banks can absorb the assets and customers of any failing bank, then only the aggregate capital of the sector matters. If the social cost of mergers is too high, however, bailouts become more attractive, which spurs moral hazard.

We next study a model where banks are imperfect substitutes, for instance because of soft information, specialization across activities and locations, or market power. Lack of substitution worsens the time-inconsistency problem as each individual bank knows it will be partly insured against its own poor returns to the extent that it would be costly for other banks to pick up the slack. We introduce the concept of ϵ-commitment to ensure continuity of the limit of mechanisms. A mechanism is ϵ-credible if welfare deviates by less than ϵ from its ex post optimum. We can then recast our first result in more general terms. We show that the 'size' of the set of implementable outcomes is proportional to $\epsilon\eta$ where η is the elasticity of substitution between capital surpluses located in different banks. The Acharya et al. (2016) loss function assume $\eta = \infty$ which is why the first best is always implementable without any commitment. On the other hand, when η is small, the first best is not implementable in the usual (strong) time consistent fashion.

When banks are differentiated, however, the notion of renegotiation-proof mechanisms in Fudenberg and Tirole (1990) becomes quite appealing. If the government promises a set of transfers, banks can block a deviation that would leave them worse off. Under this weaker form of time consistency the government can choose among Pareto optimal allocations. The government cannot directly commit to punish weak banks but it can commit not to renege on its promised support to well-performing banks. The core time-inconsistency problem is still present but our tournaments can once again implement the first best level of safety, albeit at a higher cost (that is, larger bailouts) than in the case of perfect bank substitutability. Numerically, we find that the cost decreases rapidly towards the first best cost as banks become more substitutable.

Finally, we consider a different form of heterogeneity, arising from financial linkages between banks that generate comovement in returns. These linkages capture a variety of "contagion" forces, such as cross-exposures, fire sales, or domino effects, as studied in the financial networks literature. We show how contagion leads to a natural notion of systemic risk: banks are more systemic when their performance has a stronger effect on the rest of the system. In turn, more systemic banks should act more prudently, and so a resolution mechanism must strive to give them stronger incentives. Ex post, however, the government may consider highly systemic banks "too interconnected to fail" (Haldane, 2013). Our main finding is that the constraints that financial linkages impose on bank resolution depend crucially on how bailout funds attributed to one bank spill over to other banks.

If a form of "ring-fencing" or earmarking applies to public funds and bailout money cannot flow throughout the system to benefit other banks indirectly, our tournament mechanism remains credible and efficient under minor amendments. A bank's rank in the tournament is determined by its ex post performance, as in the baseline model, but now weighted by its systemic risk. On the other hand, moral hazard comes back when earmarking public funds is not possible. Spillovers reduce ex post intervention costs to the extent that injecting money in one bank can also stabilize other banks. The problem, however, is that spillovers make it ex post optimal to always save the most systemic bank first. That systemic bank is completely insured and thus maximizes its risk-taking. Our model thus shows the importance of earmarking public funds and of limiting safe harbor provisions for interbank liabilities.

Related literature Bailouts are risky bets. Some succeed, some drag down the sovereign, as shown in Acharya et al. (2014). There is ample theoretical and empirical support for the idea that the expectation of bailouts distort incentives and create moral hazard. Kelly et al. (2016) show that the key factor affecting the pricing of financial crash insurance is the extent of collective government guarantees. Dam and Koetter (2012) find that a change of bailout expectations by two standard deviations increases the probability of official distress.

Our main contribution is to show how to use the classic rank-order tournament mechanisms of Lazear and Rosen (1981) to overcome the pervasive time inconsistency problem that generates or worsens moral hazard in bank risk-taking (Farhi and Tirole 2012, Keister 2016, Chari and Kehoe 2016).

Our results differ from existing results in the literature in two important ways. The first difference centers around commitment and tournaments. Chari and Kehoe (2016) study an economy where a utilitarian planner distorts an ex post allocation which is otherwise a Pareto optimum. Chari and Kehoe (2016) thus assume an extreme form of lack of commitment which would be solved by a renegotiation-proof mechanism (Fudenberg and Tirole, 1990). Farhi and Tirole (2012), on the other hand, study a model with symmetric banks and consider only symmetric contracts, which rule out tournament incentives.

Second, the literature argues that the moral hazard problem is worst in countries with ample fiscal space: the narrative is that if banks expect the sovereign to be able to bail them out even in deep crises, they have no reason to self-insure. We find that fiscal capacity has the opposite effect once richer mechanisms such as ours are used. Since a sovereign with larger fiscal capacity is able to transfer a larger amount to the banking sector as a whole, it also has more flexibility in the distribution of transfers across banks, which tends to relax incentive constraints and reduce moral hazard.

Keister and Mitkov (2021), Dewatripont and Tirole (2018), and Clayton and Schaab (2021) study the design of bail-in policies; we simplify the capital structure side by considering only two classes of liabilities, hard deposits and "total loss absorbing capacity" including equity and bailinable debt. Our extension to financial contagion relates to the work of Demange (2020) on resolution among interconnected banks. Our paper also relates to the strategic substitutability among banks during ex post fire sales, and the resulting ex ante incentives to build financial resilience, as in Perotti and Suarez (2002), Acharya and Yorulmazer (2007), or Malherbe (2014). Instead of considering strategic substitutability driven by a competition for cheap assets, we show how a well-designed competition for government support can implement efficient ex ante safety. Acharya and Yorulmazer (2008) also show that liquidity support to surviving banks instead of failed ones improves banks' incentives to differentiate their exposures rather than to herd. Our approach relates to Kasa and Spiegel (2008), who show that using relative instead of absolute performance evaluation in bank closures can reduce costs. Unlike us, they do not consider how a tournament-like mechanism can implement the first best risk-taking. They also assume that regulators can fully commit, while our core insight is that tournaments mitigate the time-consistency problem.

We abstract from the dynamic dimension of crises, but uncertainty and learning would only reinforce our results. Nosal and Ordonez (2016) show that uncertainty about the severity of the crisis can prompt governments to delay bailouts until it becomes clear that the crisis is systemic. This in turn gives banks incentives to make sure they survive until the government intervenes. Instead of focusing on how exogenous uncertainty improves incentives, we show that even in a perfectly known systemic crisis—hence even when bailouts are inevitable—the government can still optimally *design* asymmetric transfers to reach the first best safety.

1 A Model of Systemic Crises And Government Interventions

We present our baseline environment before defining the first best allocation. The key feature of our model is that banks decide how much risk to take, anticipating government support policies in case of a systemic crisis that hits many banks at the same time.

1.1 Environment

We consider a two-period model with $N \geq 2$ banks and a "government", that should be viewed as combining fiscal and monetary authorities. At $t = 0$, the government announces a bailout rule mapping realized returns on banks' assets to government transfers, as described below. Each bank then chooses and safety investment $x_i \in [0, 1]$. Uncertainty is resolved at time $t = 1$. Uncertainty consists of aggregate as well as bank specific shocks. We define state $s = 0$ as the normal state and the states $s \neq 0$ as the crisis states. The probability of the normal state is $\mathbb{P}[s = 0] = p_0$. The crisis states are distributed on some compact set $\mathcal{S}$ so that $\int_{\mathcal{S}} p_s ds = 1 - p_0$.

Banks. At time 0 banks have assets a and deposits with face value d due at time 1. We denote by r_i^s the gross asset return of bank i in state s at time 1 and by $m_{i,s}$ the cash injection from the government. Table 1 shows the balance sheet of bank i at time 1.

Table 1: Balance Sheet

Assets	a_i	Liabilities	
Gross Value	$r_i a_i$	TLAC	e_i
		Deposits	d_i

TLAC means total loss absorbing capacity and denotes the sum of equity (tier 1) and other loss absorbing capacity such as junior unsecured bailinable bonds. Our model has nothing new to say about of ex ante capital requirements or differences in asset liquidity. We therefore lump the various layers of TLAC into one category that we call capital, and we lump all assets returns into one category that we call gross value, or output.[1]

[1] Keister and Mitkov (2021) study the interaction between private incentives to bail in investors and public incentives to bail them out. Similarly, Dewatripont and Tirole (2018) endogenize the composition of liquid and illiquid assets.

We say that a bank is well capitalized when $e_i \geq \underline{e} a_i$ or equivalently $r_i \geq \underline{r}_i = d_i/a_i + \underline{e}$, and its capital surplus is then $e_i - \underline{e} a_i$. This notion of well capitalized is defined in the welfare function below. Banks maximize expected capital returns net of transfers. The gross returns are given by

$$r_i^s = \begin{cases} f(x_i) + \xi_i & \text{with probability } p_0 \\ r_{i,s} \sim G(. \mid x_i, s) & \text{with probability } p_s \end{cases} \tag{1}$$

The shocks ξ_i are positive (hence $f(x_i)$ is the minimum gross return in normal times) and i.i.d. across banks and the crisis returns $r_{i,s}$ are bounded. The expected return in the normal state f is decreasing, bounded, and concave over $[0, 1]$ and attains a strict maximum at 0. The shock s is common to all banks. The cumulative distribution $G(x_i, s)$ of the return $r_{i,s}$ is ranked by stochastic dominance.[2]

Assumption 1. $G(r \mid x_i, s)$ *is decreasing and continuously differentiable in* x *for all* r.

The function f thus captures the risk/return tradeoff that banks face. Banks can improve their crisis return by increasing x, at the cost of lower returns $f(x)$ in normal times. The maximal risk banks can take, $x = 0$, leads to a highest expected return $f(0)$ in the good state but the worst exposure in crisis states.

Government. The government observes the aggregate state at time 1 as well as the banks' returns $r_{i,s}$. We will normalize the parameters of the model so that the normal state is indeed normal, i.e, no crisis and no bailout. The government's ex post value

$$V\left(\{e_i, a_i\}_{i=1..N}\right)$$

is concave and weakly increasing in e_i, decreasing in a_i. To simplify the notations we often write $V\{e_i\}$ since $\{e_i\}_{i=1..N}$ are the only random parts of the function, but the function itself also depends on a_i and the parameter $\underline{e}$. Finally, V is flat at its maximum when all banks are well capitalized: $V = \bar{V}$ when $e_i \geq \underline{e} a_i$ for all $i = 1..N$. This *defines* what we mean by a "well capitalized" banking system. Our formulation based on a general value function V encompasses multiple (and non-exclusive) frictions that

[2] In Section 6 we will allow the distribution of $r_{i,s}$ to depend on other banks' safety investments x_j as well.

arise when bank capital is low, even when banks are still solvent. We discuss micro-foundations for V below in terms of runs and credit crunch.

The government has the option to mitigate the consequences of financial distress by implementing transfers $\{m_{i,s}\}$. The total cost $M = \sum_i m_{i,s}$ is subject to a shadow cost of public transfers $\Gamma(M;\gamma)$ which is positive, weakly convex and strictly increasing for all $M > 0$. We index the cost of funds to $\gamma \geq 0$ which measures the inverse of fiscal slack. The function $\Gamma(M;\gamma)$ is increasing in γ and super-modular in (M,γ). Ex ante aggregate welfare is thus defined as

$$\mathbb{E}\left[R + V\{e_{i,s} + m_{i,s}\} - \Gamma(M_s;\gamma)\right]. \tag{2}$$

where $R = \sum_i a_i r_{i,s}$ is the random aggregate asset return.

Discussion of Assumptions The results of the paper do not depend on the specific friction that gives rise to the welfare value V, but for concreteness we provide examples of micro-foundations in Appendix A. Broadly speaking, two classes of models can deliver the welfare function specified above. The first class includes models of runs (Diamond and Rajan, 2012). A bank with low equity (but still potentially solvent) may face the risk of a run, unless it restructures part of its debt; restructuring, however, can trigger money market disturbances (further runs, as we saw after the collapse of Lehman Brothers).

The second class includes models of credit crunch (Myers, 1977; Holmström and Tirole, 1997; Philippon and Schnabl, 2013). In these models, a new investment opportunities arises at date-1, but limited pledgeability (or other frictions such as debt overhang) prevents solvent banks from investing at the efficient scale unless they bring enough equity/liquidity into the period. The welfare cost in models of runs comes from fire sales (Stein, 2012) or from the inefficient liquidation of existing assets. In models of credit crunch the welfare cost arises from inefficiently low investment in new projects. Both costs are clearly relevant and the Appendix shows how each maps into a welfare function V.[3]

We wish to focus our analysis on the issue of systemic risk, not on the pricing of deposit insurance. We therefore assume

[3]One advantage of using a welfare function V is to highlight the key feature that is not typically discussed in micro-founded models. As our analysis makes clear, the critical feature is the substitutability of capital between banks with a shortfall and banks with a surplus. In a credit crunch model, then, the key feature is whether bank 1 can lend to the customers of bank 2, either directly or after a merger when bank 2 is distressed. Standard models of runs, fire sales and credit crunch typically do not highlight this aspect.

Assumption 2. ***Well calibrated TLAC.*** $\frac{d_i}{a_i} \leq \min\{r_{i,s}\} < \underline{r}_i = \frac{d_i}{a_i} + \underline{e}$.

Assumption A2 means that TLAC requirements are calibrated so as to protect small depositors, and allows us to focus on the issue of undercapitalization during systemic crises.

The variable x captures the efforts of the bank to mitigate its systematic risk. It includes investment in liquid or safe assets with a low return as well as investments in monitoring and screening technologies and risk governance in general. We assume that x is not contractible. More precisely, we think of x as the residual discretion that bankers have once they have fulfilled their quantitative regulatory requirements, such as Tier 1 ratios, TLAC and LCR. The post crisis policy response has focused on ensuring a minimum level x but these regulations are necessarily imperfect due to informational delays, signal jamming, off-balance sheet transactions, etc. Some private sector discretion always remains, so we normalize the regulatory level of safe investment to zero and view x as the residual investment in safety, above and beyond what can be enforced ex ante. Our baseline model ignores direct contagion between banks. We extend the model to allow for contagion in Section 6.

The variable m_i is the net transfer to bank i across all discretionary policies: the most obvious interpretation is that of direct equity injections, but we can also think of other implicit and explicit subsidies such as credit guarantees and loans at a reduced interest rate. Philippon and Skreta (2012) and Tirole (2012) discuss these policies in the context of an adverse selection model, and Diamond and Rajan (2011) and Philippon and Schnabl (2013) in the context of a debt-overhang model. What matters in our model is the net subsidy component of these policies, i.e., the excess payment that the government makes compared to current market prices.

Finally, our paper focuses on payoffs in the crisis state. In general, the planner might want to use information from the normal state to provide ex ante incentives. In practice there are two reasons why this is not feasible. The empirical reason is that returns in normal states contain little information about returns in crisis states. For instance, Acharya et al. (2016) find that the cross-section of returns only begin to predict returns during the GFC after the end of 2006. Relative returns during the boom years contain no useable information for estimating performance during the crisis. We thus assume that $\mathbb{VAR}(\xi_i) \gg \mathbb{VAR}(\epsilon_i)$. The theoretical reason is that $f(x_i)$ is a decreasing function of x so an incentive scheme would have to punish a firm for good performance and these schemes are not robust to hidden trading as shown in Innes (1990) and Nachman and

Noe (1994).

1.2 No Bailout

Consider first the allocations when bailouts are ruled out by assumption.

We start with the privately optimal solution. Under A2, maximizing e_i is equivalent to maximizing $r_{i,s}a_i$. Let $\tilde{x}$ be the privately optimal safe return of a bank anticipating $m = 0$ in all states:

$$\tilde{x}_i \equiv \arg\max_{0 \leq x_i \leq 1} p_0 a_i f(x_i) + (1 - p_0) a_i \mathbb{E}[r_{i,s} \mid x_i]. \tag{3}$$

By stochastic dominance the function $\mathbb{E}[r_{i,s} \mid x]$ is increasing in x and the concavity of f guarantees the existence of a unique solution. Since have assumed that the safety investment set is the same for all banks, $\tilde{x}_i = \tilde{x}$ is the same for all i.

Consider next the socially optimal allocation. Since f is concave it is optimal for the planner to set the same level of safety for all the banks. The return in the normal state is therefore $\sum_i f(x_i)$ and $\sum_i r_{i,s}$ in a crisis state. We can define the no-bailout optimal solution as

$$\mathbf{x}_0^* = \arg\max_{\mathbf{x}} \sum_i a_i \left(p_0 f(x_i) + (1 - p_0) \mathbb{E}[r_{i,s} \mid x_i]\right) + \mathbb{E}\left[V\left(\{e_{i,s}\}_i\right) \mid \mathbf{x}\right] \tag{4}$$

where $\mathbf{x}_0^* = \left(x_{1,0}^*, .., x_{N,0}^*\right)$ is the vector of safety investment by banks. The concavity of V guarantees the existence of a unique solution. We maintain throughout the paper the assumption that banks are well capitalized in the normal state. We also assume that the efficient safety investment without bailout is positive.

Assumption 3. *$0 < x_{i,0}^*$ and $f\left(x_{i,0}^*\right) > \underline{r}_i$ for all i.*

Note that, since V is an increasing function, we have $x_{i,0}^* \geq \tilde{x}$ for all i. Even without bailouts, the planner prefers higher safety investments than what banks would choose individually due to the externality captured by V.

1.3 First Best Allocation with Bailouts

Define $M \equiv \sum_i m_i$ as the state contingent aggregate bailout. Assumption A2 guarantees that $M = 0$ in the normal state since the option to bailout can only decrease the optimal level of ex ante safety (i.e., the solution of the full program is always such that $x^* \leq x_0^*$, therefore $f(x^*) > \underline{r}$ since f is decreasing). The program of the planner is therefore

$$(\mathbf{x}^*, \mathbf{m}^*) = \arg\max_{\mathbf{x},\mathbf{m}} p_0 \sum_i a_i f(x_i) + (1 - p_0) \sum_i a_i \mathbb{E}\left[r_{i,s} \mid x_i\right] \\ + \mathbb{E}\left[V\left(\{r_{i,s}a_i + m_{i,s} - d_i\}_i\right) - \Gamma(M;\gamma) \,|\mathbf{x}\right]$$

We define the ex post optimal vector of bailouts as

$$\mathbf{m}^*(\mathbf{r}) \equiv \arg\max_{\{m_i\}_i} V\left(\{r_{i,s}a_i + m_{i,s} - d_i\}_i\right) - \Gamma(M;\gamma).$$

A positive bailout in the worst state is typically part of the first best allocation. This is in line, for instance, with the theoretical results in Keister (2016) in the context of a Diamond and Dybvig (1983) model. More generally, it is not difficult to imagine that the government is more efficient than the private sector at providing some form of catastrophe insurance. In this case, it would be inefficient to force the private sector to fully self-insure against very unlikely but costly crises. The issue is therefore not the existence of strictly positive bailout probability, but rather what the anticipation of a bailout does to private incentives for safety.

2 A Pure Systemic Risk Model

In this section we follow Acharya et al. (2016) and assume that the value function depends only on the *aggregate* capital surplus of the banking sector:

$$V(\{e_i, a_i\}) = V\left(\sum_i (e_i - \underline{e}a_i)\right) \tag{5}$$

where V is increasing and concave. For instance, the systemic expected shortfall in Acharya et al. (2016) uses the piecewise linear case $V = \min\{0, \sum_i (e_i - \underline{e}a_i)\}$. The assumption behind this loss function is that the banking sector has specific expertise that is not easily replicated by non-bank actors, but that banks within the sector are good substitutes for one another. With this loss function, the government does not care about the distribution of returns across banks, but only about the aggregate capital shortfall of the banking sector. In other words, we assume that the expertise that makes banks socially valuable, for instance their ability to lend to SMEs and households, is transferable across banks but not outside the banking system. If a bank fails, its outstanding assets and new lending can be picked up by other surviving banks. By definition, when the system is solvent, it is possible to transfer assets and liabilities to solvent banks. By contrast, when the banking system is insolvent, the planner cannot avoid a disruption that has real welfare costs because it is costly to transfer bank assets outside the banking sector, either to deep-pocket private investors or to the government itself, and it is difficult to raise bank equity quickly in a crisis.

We relax these assumptions in later sections, but view them as a good starting point to capture the deadweight loss from an undercapitalized banking system. In Section 4 we allow for mergers between banks and show that the pure systemic risk model of equation (5) corresponds to the case of low merger costs, in the sense that activities of distressed banks can be transferred to strong banks with limited disruption. In Section 5 we consider imperfect substitutability between banks: individual shortfalls matter because some institutions are "too-specific-to-fail". We generalize our results to show that the amount of commitment needed to implement the first best is inversely proportional to the degree of substitutability between banks.

2.1 Ex Post Optimal Bailout

Define the *aggregate* return as $R \equiv \sum_i a_i r_{i,s}$ and the aggregate gross requirement as $K \equiv \sum_i (\underline{e} a_i + d_i)$. The ex post optimal bailout is then simply a function of the aggregate return. We define the maximized value function as

$$\mathcal{V}(R-K;\gamma) \equiv \max_{M \geq 0} V(R+M-K) - \Gamma(M;\gamma),$$

and the optimal bailout as

$$\mathcal{M}(K-R;\gamma) \equiv \arg\max_{M \geq 0} V(R+M-K) - \Gamma(M;\gamma). \tag{6}$$

Proposition 1. *The maximized value function $\mathcal{V}$ is increasing and concave in $R-K$, and decreasing in γ. The bailout $\mathcal{M}(K-R;\gamma)$ is increasing in $K-R$ and decreasing in γ. There exists a threshold $\mathcal{K}(\gamma) \in [0,K]$, decreasing in γ such, that $\mathcal{M}=0$ for $R \geq \mathcal{K}(\gamma)$.*

The value function $\mathcal{V}$ is concave and differentiable irrespective of the shape of V and Γ. The bailout function, on the other hand, may or may not be convex, and is usually not differentiable. For instance, when the systemic externality is piecewise linear $V = \min(0, E - \underline{e}A)$ and the fiscal cost of funds is quadratic $\Gamma = \gamma M^2$, then the bailout is flat at $(2\gamma)^{-1}$ when the crisis is severe and then linearly decreasing (in R) to zero when the return is between $K - (2\gamma)^{-1}$ and K.

Example: Linear Cost of Funds Suppose that the cost of funds is linear

$$\Gamma(M) = \gamma |M|$$

The quasi-linear preferences of the planner imply that the ex post optimal bailout takes the simple form of a put option on the aggregate return R:

Lemma 1. *With linear cost of funds, the optimal aggregate bailout is*

$$\mathcal{M} = \max\{0, \mathcal{K}(\gamma) - R\}$$

where $\mathcal{K}(\gamma) \in [0, K]$ is decreasing.

The planner has an aggregate target $\mathcal{K}(\gamma)$ which depends on the aggregate capital requirement K and the cost of public funds γ. If the private sector delivers the target by itself ($R > \mathcal{K}$), then the planner does not intervene. If the private sector falls short of the target ($R < \mathcal{K}$) then the planner replenishes aggregate capital up to the target to $\mathcal{M}(R) + R = \mathcal{K}$. The replenishment may not be complete ($\mathcal{K} < K$) when public funds are costly and when V approaches its maximum smoothly from the left.

2.2 First Best

With the welfare function (5), the first best solution solves

$$\mathbf{x}^* = \arg\max_{\mathbf{x}\geq 0} p_0 \sum_i a_i f(x_i) + (1-p_0)\sum_i a_i \mathbb{E}[r_{i,s} \mid x_i] + \mathbb{E}\left[\mathcal{V}\left(\sum_i a_i r_{i,s} - K\right) \mid \mathbf{x}\right].$$

The loss function is decreasing in R and increasing in γ which implies that

$$\tilde{x} \leq x_i^* \leq x_{i,0}^*.$$

The planner always wants more safety than the privately optimal choice under no bailout $\tilde{x}$, but requires less than in the optimal case without bailouts x_0^* because the option to bail out limits downside risks.

Notice that optimal safety may depend on bank size because of the non-linear loss function.

Lemma 2. *Let $G_\epsilon(. \mid x_i, s)$ be the distribution of $\epsilon_i = r_{i,s} - \mathbb{E}[r_{i,s} \mid x_i, s]$ and let $\varepsilon \equiv \sum_i a_i \epsilon_i$ be the aggregate of bank-level shocks. Optimal safety does not depend on size when G_ϵ does not depend on x.*

We get scale independence if return volatility does not depend on x. An example is $r_{i,s} = \alpha(x_i) + s + \epsilon_i$ where α is increasing. This implies $R = \sum_i a_i \alpha(x_i) + As + \varepsilon$ where ε is independent of $\mathbf{x}$. On the other hand there are realistic cases where x would affect the volatility of r. For instance, if $r_{i,s} = \alpha(x_i) + s + (1 - x_i)\epsilon_i$, efficiency requires large banks to invest more in safety.

We say that a crisis is systemic if it necessitates a bailout (i.e., when $R < \mathcal{K}$) and moderate otherwise. We summarize our results in the following proposition.

Proposition 2. *The social optimum is characterized by $(\mathbf{x}^*, \mathcal{M}(K - R; \gamma))$. Safety investments $\mathbf{x}^*$ are increasing in γ, in aggregate banking assets A, and in the mean and variance of s; they are decreasing in $\underline{e}$ and satisfy $(\tilde{x}, ..\tilde{x}) \leq \mathbf{x}^* \leq \mathbf{x}_0^*$.*

Propositions 1 and 2 put some discipline on the range of outcomes that are consistent with optimal regulations and interventions. There are no bailouts in moderate states. Once the capital shortfall is large enough, the planner finds it optimal to transfer bailout funds to banks. The shape of the bailout is then pinned down by fiscal capacity. When the fiscal cost is linear (e.g., the US), it is optimal to fully insure the banking system against further downside risk. When the fiscal cost is convex (e.g., Ireland, Greece, Cyprus), the bailout increases less than one for one with the losses.

In the first best, the government *mandates* the optimal safety vector $\mathbf{x}^*$, thus avoiding moral hazard. In the rest of the paper we study what happens when x is unobserved by the government. The model then includes the potential for a strong form of moral hazard. When $M^* > 0$ the aggregate return net of government transfer does not depend on x. Anticipating this, banks might discount the systemic states and increase their risk taking.

2.3 Moral Hazard under No Commitment and Symmetric Bailouts

We now assume that x cannot be observed and we impose a time-consistency, or "credibility", constraint. The government is restricted to rules $\{m_i\}$ that are ex post optimal, even off the equilibrium path. Therefore

$$\sum_i m_{i,s} = \mathcal{M}\left(K - R\right) \tag{7}$$

for all possible values of R where $\mathcal{M}\left(K - R\right)$ is defined in (6). We define a symmetric bailout as follows.

Definition 1. A bailout is symmetric if, for all $(i, j) \in [1 : N]^2$ and all $s \in \mathcal{S}$, we have $\frac{m_{i,s}}{a_i} = \frac{m_{j,s}}{a_j}$.

When all banks of ex ante identical a symmetric bailout is one where they all get the same amount of money. When banks' sizes vary, the definition simply allows proportionality with size. In a symmetric bailout satisfying the credibility constraint (7) we must have $m_{i,s} = a_i \frac{\mathcal{M}(R)}{A}$. The best response of bank i is therefore

$$\beta_i\left(\mathbf{x}_{-i}\right) = \arg\max_{x_i \geq 0} p_0 a_i f\left(x_i\right) + \left(1 - p_0\right) a_i \left(\mathbb{E}\left[r_{i,s} \mid x_i\right] + \Omega\left(x_i; \mathbf{x}_{-i}\right)\right) \tag{8}$$

where $\mathbf{x}_{-i}$ is the vector of safety investments by all banks except bank i, and Ω is defined as $\Omega\left(\mathbf{x}\right) \equiv \frac{1}{A}\mathbb{E}\left[\mathcal{M}\left(K - R\right) \mid \mathbf{x}\right]$, which we can write as

$$\Omega\left(\mathbf{x}\right) = \frac{1}{A}\int \mathcal{M}\left(K - R\right) d\Phi_N\left(R \mid \mathbf{x}\right). \tag{9}$$

The distribution Φ_N is the convolution of the underlying ones: $R \mid \mathbf{x} \sim \sum_{i=1}^{N} a_i r_{i,s} \mid \mathbf{x}$.

Lemma 3. $\Omega\left(\mathbf{x}\right)$ *is continuous, decreasing in each* x_i*, and satisfies the increasing differences condition in* $\left(x_i, \mathbf{x}_{-i}\right)$ *for all* i*.*

Lemma 3 immediately implies that, for all possible values of $\mathbf{x}_{-i}$, the best response is bounded above by the private equilibrium: $\beta\left(x_{-i}\right) \leq \tilde{x}$. Our game takes place on compact sets with a finite number of players, continuous choices and continuous reward functions, therefore we know that at least one Nash equilibrium exists and any solution satisfies $\hat{x} \leq \tilde{x}$. We summarize our discussion in the following proposition.[4]

[4]Given risk-neutrality, it is without loss of generality to focus on pure strategies. Fudenberg and Tirole (1990) show that with risk-averse agents, it is possible to maintain some incentives once we allow for mixed strategies.

Proposition 3. *All equilibria with no commitment and symmetric bailouts have the following properties:*
(i) Lack of commitment creates strategic complementarities in risk taking: $\beta_i(\mathbf{x}_{-i})$ *is increasing.*
(ii) Safety is too low ($\hat{x}_i < x_i^*$) *and the probability of a systemic crisis is too high:* $\Phi_N(K \mid \hat{\mathbf{x}}) > \Phi_N(K \mid \mathbf{x}^*)$.
(iii) Safety decreases when the cost of public funds γ *decreases.*
(iv) If $\beta_i(\mathbf{0}) = 0$ *a full unraveling equilibrium exists with minimum safety, maximum systemic risk, and maximum bailout* $x_i = 0$ *for all* i.

Lack of government commitment creates strategic complementarities between banks: if all banks reduce their safety the probability of a bailout increases, which reduces the marginal incentives to hedge against systemic crises. Lack of government commitment can generate an extreme form of moral hazard where banks make no investment in safety. A marginal increase Δx_i reduces the bank's expected bailout. We have illustrated this point in the simple case of symmetric bailouts, but more generally it will hold whenever the expected bailout $\mathbb{E}[m_i|\mathbf{x}]$ received by bank i is decreasing in its own safety x_i.

Strategic Complementarities and Uniqueness While strategic complementarities are a realistic feature, they can open the door to multiple equilibria if those complementarities are too strong. We can in principle deal with multiple equilibria: there is a set of equilibria, and each time we say that safety is increasing we mean it in the Strong Set Order sense of Topkis (1978) and Milgrom and Shannon (1994). Alternatively, we could allow the government to act as a coordination device and select the equilibrium with highest safety. These solutions are feasible but they create a large burden of notations without changing the economic insights. It is more convenient to have a unique equilibrium to state our main results in the next section. We therefore assume that Ω is not too convex or that f is concave enough.

Assumption 4. *The slope of the best response* $\beta_i(\mathbf{x}_{-i})$ *is less than one.*

3 Credible Tournaments

The previous section has shown that when the government lacks commitment, standard bailout mechanisms lead to moral hazard. In stark contrast, we now show that the government can use relative performance evaluation among multiple banks to solve the moral hazard problem and implement the first best allocation in a time-consistent fashion. The reason is that the credibility constraint only affects the aggregate bailout, and leaves enough leeway to the government to structure the *distribution* of bailouts across banks. In particular, the government can use a relatively simple tournament scheme that rewards banks according to their ranking while maintaining credibility. For simplicity we illustrate our main result in the case where banks are ex ante identical, thus assuming $a_i = 1$ for all banks; we extend our mechanism to account for heterogeneous bank size in Appendix B.

3.1 Bonus-Malus Implementation

Two Banks. We build intuition by considering the case of two banks. We define the tournament rule $\mathscr{T}$ with two banks as

$$m_i = \begin{cases} \frac{\mathcal{M}(K-R)}{2} + \Delta & r_{i,s} > r_{j,s} \\ \frac{\mathcal{M}(K-R)}{2} - \Delta & r_{i,s} < r_{j,s} \end{cases}$$

Note that $\mathbb{P}[r_{1,s} > r_{2,s} | \mathbf{x}] = H_s(x_1, x_2)$ where H_s is increasing in x_1 and decreasing in x_2. The best response function for bank 1 is therefore

$$\hat{x}_1 = \beta_1(\Delta, x_2) = \arg\max_{x_1} p_0 f(x_1) + (1-p_0)(\mathbb{E}[r_{1,s} \mid x_1] + \Omega(x_1, x_2)) \\ + 2\Delta \int_s H_s(x_1, x_2) p_s ds.$$

The crucial departure from perfect insurance and the ensuing moral hazard comes from Δ, which rewards the best bank and punishes the other one. When $\Delta = 0$ this best response corresponds to the one discussed in Proposition 3. We can then state our first main proposition.

Proposition 4. *With $N = 2$, there exists a unique $\Delta^* > 0$ that implements the social optimum $(x^*, x^*, \mathcal{M}(K-R))$.*

Note that Δ^* is unique in the class of mechanisms that we consider but there are other classes of mechanisms that can implement the first best. We know from Proposition 3, however, that all of them must use some form of relative performance evaluation.

N **Banks.** It is straightforward to extend our results to N banks. In fact, it is easier than with two banks since there are more degrees of freedom. A possible rule is

$$m_i = \frac{\mathcal{M}(K-R)}{N} + \Delta \times \mathcal{I}(r_i - \mathrm{med}(\mathbf{r}))$$

where the function $\mathcal{I}$ is such that $\mathcal{I}(y<0) = -1$, $\mathcal{I}(0) = 1$, and $\mathcal{I}(y>0) = 1$ and $\mathrm{med}(\mathbf{r})$ is the median return. By definition of the median

$$\sum_i^N \mathcal{I}(r_i - \mathrm{med}(\mathbf{r})) = 0$$

so $\sum_i^N m_i = \mathcal{M}(R)$ and the rule is credible. Denote $H_{s,N}^{\mathrm{med}}(x_i, x_{-i})$ the probability that $r_i > \mathrm{med}(\mathbf{r})$ when other banks play $\mathbf{x}_{-i}$ and bank i plays x_i. $H_{s,N}^{\mathrm{med}}$ is increasing in x_i and decreasing in $\mathbf{x}_{-i}$. Then bank i solves

$$\hat{x}_i = \beta_i(\Delta, \mathbf{x}_{-i}) = \arg\max_\theta (1-p_0)\, x_i + (1-p_0)\left(\mathbb{E}[r_{i,s} \mid x_i] + \Omega(x_i, \mathbf{x}_{-i})\right) + 2\Delta \int_s H_{s,N}^{\mathrm{med}}(x_i, \mathbf{x}_{-i})\, p_s ds.$$

Following the same steps as for $N = 2$ we have:

Proposition 5. *For any number $N \geq 2$ of banks, there exists a unique $\Delta^* > 0$ that implements the social optimum $(\mathbf{x}^*, \mathcal{M}(K-R))$.*

The simplicity of our "median" rule makes it attractive, but there are many other more complex rules that can achieve the same objective, even within the class of tournaments. For instance, different prizes could be attributed to banks according to their exact ranking in terms of returns, and not just whether they are above or below the median.

The implementation above might require large punishments in equilibrium. A bank with a bad draw needs to be punished to provide ex ante incentives. There are, however, practical limits on punishments. The first limit, which we consider next, is that the planner might not be *able* to punish because of limited liability. The second limit, which we study in Section 5, is that the planner might not be *willing* to punish because of imperfect substitutability between banks.

3.2 Limited Liability

Let us now consider the case where government transfers and taxes are constrained by limited liability (LL). There are two ways to write limited liability. The strict form ("strict LL") is $m_i \geq 0$ for all banks in all states, which simply rules out negative transfers. This constraint typically leaves equity holders with a surplus. A weaker form of limited liability ("weak LL") is $a_i r_i + m_i \geq d_i$, which allows negative transfers of residual equity value, but not more. In Section 3.3 we show how these two cases can be interpreted as polar cases of a richer model with fire sales and mark-to-market accounting in resolution. Our general result holds under strict (and therefore also weak) limited liability.

Proposition 6. *Even under strict limited liability* ($m_i \geq 0$)*, tournament incentives rule out moral hazard* ($\hat{x} > \tilde{x}$) *and implement the first best when the cost of funds is low, i.e., there exists* $\hat{\gamma} > 0$ *such that* $\hat{x} = x^*$ *for any* $\gamma < \hat{\gamma}$*.*

What happens when the cost of funds γ is above the threshold $\hat{\gamma}$, making the first best unattainable? Characterizing the second best allocation is a lot more complicated, so we use the following special case with binary outcomes. We assume that all banks are ex ante identical with size a. At time 1 banks are randomly allocated into two groups, L and H, with sizes N_L and N_H such that $N_L + N_H = N$. The returns of bank i are determined jointly by its risk management, its group, and the aggregate state:

$$r_i^s = \begin{cases} f(x_i) + \xi_i & \text{in the normal state} \\ s + x_i \mathbb{I}_{i \in H} & \text{in crisis state } s \end{cases}$$

The model thus works as follows. Banks make ex ante safety choices x_i. If a bank ends up in group L its return is s irrespective to x. If a bank is in group H, its safety choice matters ex post as its return is $s + x_i$. The key point is that there is no way for the planner to distinguish a bank in group L from a bank in group H who chose $x = 0$. In this setting we obtain the following result:

Proposition 7. *The highest implementable safety under limited liability is decreasing in the cost of public funds* γ *and decreasing in the size of the banking sector* A*. Incentives and ex ante leverage constraints* d/a *are substitutes under strict LL, but complement under weak LL.*

Proposition 7 gives a striking result with respect to fiscal slack: a lower γ *increases* safety. This is exactly the opposite of the conventional wisdom based on symmetric mechanisms. The result under weak LL also gives a novel rationale for leverage limits or higher capital requirements; we generalize the result and provide more intuition in the next subsection on fire sales. It also gives a macro-prudential reason for clawback provisions on executive compensation to reduce the binding limited liability constraint.

One should keep in mind that taxes can also be levied ex ante, for instance to provision a "bailout insurance fund". Banks could all pay the same tax at time 0 and recoup different payments at time 1 based on the tournament rule. This would improve incentives by effectively relaxing the limited liability constraint.

3.3 Fire Sales

We conclude this section by generalizing the polar cases of weak and strict LL in a simple model of fire sales. Under our scheme, fire sales are useful for incentives, because lower secondary market prices decrease the outside option of the distressed banks which helps relax their limited liability constraint. Moreover, incentives and ex ante leverage constraints are complement if and only if fire sale prices are low enough.

Suppose that during the crisis, the regulator is constrained to net transfers m_i that cannot expropriate bank shareholders *at current market prices*. Thus shareholders have the choice between accepting resolution and obtaining a payoff $ar_i + m_i - d$, with assets left at book value within the bank until the crisis is over, or liquidating assets at fire sale prices immediately. We can interpret the return r_i as the fundamental value that assets recover to after the crisis. In the midst of the crisis, however, asset values can be temporarily lower, equal to $(1-\chi)r_i$, where $\chi \in [0,1)$ is a fire sale discount on assets.[5] Therefore the shareholder participation constraint is

$$\begin{cases} m_i + ar_i \geq d & \text{if } r_i \leq \frac{d}{(1-\chi)a} \\ m_i + \chi a r_i \geq 0 & \text{if } r_i \geq \frac{d}{(1-\chi)a} \end{cases} \iff m_i \geq a \max\left\{\frac{d}{a} - r_i, -\chi r_i\right\}.$$

For deep fire sale discounts $\chi \to 1$, the constraint converges to weak LL. For moderate discounts, the constraint writes $m_i + \chi a r_i \geq 0$, and strict LL corresponds to the case without fire sales $\chi = 0$. Just like weak LL is easier to satisfy than strict LL, a deeper fire sale discount χ allows the regulator to impose tougher punishments on weak banks during the crisis, and therefore relaxes the incentive constraint for all banks ex ante.

[5]We treat χ as fixed to simplify, but our results would extend to a stochastic χ that is potentially correlated with returns, as would be the case, for instance, when endogenizing asset prices using "cash-in-the-market pricing".

We can also generalize the result in Proposition 7 on the effect of leverage:

Proposition 8. *Consider the same setting as in Proposition 7 and suppose the cost of funds is linear. Let*

$$\hat{\chi} = 1 - \frac{A}{\mathcal{K}(\gamma)} \frac{d}{a}.$$

Incentives and ex ante leverage constraints on d/a are complement if the fire sale discount χ is higher than $\hat{\chi}$, and substitutes otherwise.

Tightening ex ante leverage regulation by decreasing d/a has two effects on incentives. On the one hand the aggregate bailout is lower, which undermines incentives to become a good bank, just like tighter fiscal space does. On the other hand lower leverage allows regulator to impose a harsher penalty upon poorly performing banks, especially in the case of deep fire sale discounts (high χ). If fire sales are mild (low χ), however, regulators cannot take away much from weak banks regardless of their leverage so the negative effect on incentives dominates.

Of course, fire sale prices are endogenous and a successful intervention would also reduce χ as in Philippon and Skreta (2012) and Tirole (2012). We leave that to future work.[6]

4 Mergers and Resolution Authority

We have used the benchmark loss function $V\left(\sum_i e_i - \underline{e} a_i\right)$ to establish our first main result with and without limited liability. In all these cases policy makers intervene using taxes and transfers. These instruments are used extensively in practice, but there is another tool that is used extensively and requires a modification of the baseline model: mergers of weak banks with strong ones. In this section we endogenize the distribution of assets and liabilities by giving the government a resolution authority.

Definition 2. Resolution authority is a technology with which, for any undercapitalized bank $e_i < \underline{e} a_i$ the government can write capital claims to 0 and transfer the assets and deposits to another bank or another set of banks.

[6] Fire sales also create strategic substitutability between banks even absent any bailout mechanism, as pointed out by Perotti and Suarez (2002), Acharya and Yorulmazer (2007; 2008), or Malherbe (2014). This can even lead to predatory hoarding, which in our model would take the form of over-investment in safety x in order to buy cheap assets sold by distressed banks. We focus on the case where banks under-invest in safety under laissez-faire.

We have already discussed the issue of strict versus weak limited liability so we focus here on the case of weak limited liability where regulators have the authority to wipe out investors of failing banks. It is straightforward to extend the results to the case of strict limited liability. To discuss mergers we need to specify a value function over different sets of existing banks. We consider the following value function

$$V\{e_i, a_i\} = V\left(\sum_{i=1}^{N} a_i v(y_i)\right), \tag{10}$$

where y_i is defined as the percentage surplus of bank i

$$y_i \equiv \frac{e_i}{a_i} - \underline{\mathrm{e}}.$$

The functions V and v are increasing and (weakly) concave with $v(0) \geq 0$ and $V(0^+) = \bar{V}$. For instance, in our application below we use $v(y) = \min(0, y)$, while the benchmark model of Acharya et al. (2016) corresponds to $v(y) = y$.

This value function has two key properties that make it appealing to study mergers and capital shortfalls. The first property is that it is neutral with respect to the combination of similarly capitalized banks. If $y_i = y_j = y$ then we get $a_i v(y_i) + a_j v(y_j) = (a_i + a_j) v(y)$ so nothing is gained or lost by combining two similar banks. This is clearly a desirable feature of any welfare function. The second key property is that the concavity of v around $y = 0$ captures the degree of substitution between capital shortfalls and surpluses: $v(y) = \min(0, y)$ implies zero substitution while $v(y) = y$ implies perfect substitution, with most realistic cases somewhere in between. For instance, in a fire sales model, distressed banks are forced to sell, while banks with surpluses take advantage of low prices, but they do not pick up the slack one for one. On the other hand, this value function does not capture two economic forces that may be important in some context: taste for variety and market power. We study these issues in Section 5.

Let us now define a merger allocation and its cost.

Definition 3. A merger allocation is a matrix α where $\alpha_{i,j} \in [0, a_i]$ are the assets from bank i transferred to bank j and $\sum_{j=1}^{N} \alpha_{i,j} = a_i$. The cost of the merger allocation is $\tau \sum_{i=1}^{N} (a_i - \alpha_{i,i})$.

The idea here is simple. Mergers reallocate assets and the cost of transfer is τ per unit of assets. One cost of mergers is the due diligence required to ascertain asset quality;

for instance, during the 2007-2008 crisis some mergers famously fell through (Lehman Brothers) or almost did so (Bear Sterns) in part because regulators and potential buyers did not have enough time and resources to value complex assets in the midst of the panic. Another cost of mergers comes from the non-pledgeable rents that must accrue to employees and managers of the acquired bank (as in Hart and Moore, 1995; Holmström and Tirole, 1997). Finally, mergers can impede the flow of soft information, especially if the target is a small or specialized bank (Sapienza 2002, Stein 2002).

4.1 Frictionless Mergers: An Aggregation Result

Consider for instance the sale $\alpha_{i,j}$ from bank i to bank j. The net welfare gain is then

$$V\left(..+\left(a_i-\alpha_{i,j}\right)v\left(y_i\right)+\left(\alpha_{i,j}+a_j\right)v\left(\frac{\alpha_{i,j}y_i+a_jy_j}{\alpha_{i,j}+a_j}\right)\right)-V\left(..+a_iv\left(y_i\right)+a_jv\left(y_j\right)\right)-\tau\alpha_{i,j}. \tag{11}$$

Since v is concave and V increasing we know that the first difference is positive and the question is whether it is high enough to cover the cost $\tau\alpha_{i,j}$. Would the shareholders of bank j approve the merger? Under Assumption A2 of well calibrated TLAC the value of assets exceeds that of liabilities, so the merger increases shareholder value, but the merged bank might still be undercapitalized (if $\alpha_{i,j}y_i + a_jy_j < 0$), in which case the regulator might want to provide bailout funds to bank j. We return to these issues later.

Let N_- and N_+ the sets of undercapitalized and well capitalized banks. When we study mergers the two key state variables are the mass of failed assets $A_- = \sum_{i\in N_-} a_i$ and the aggregate surplus equity $E - \underline{e}A$. We have the following aggregation result.

Proposition 9. *Let* $V^{post} = \max_\alpha V\left(\alpha, \{e_i, a_i\}\right) - \tau\sum_{i=1}^{N}\left(1-\alpha_{i,i}\right)a_i$ *be the post merger welfare value. We have*

$$V^{post} \geq V\left(Av\left(\frac{E-\underline{e}A}{A}\right)\right) - \tau A_-.$$

As $\tau \to 0$ *any value function of the type (10) converges to the value function* $V\left(Av\left(\frac{E-\underline{e}A}{A}\right)\right)$.

The proposition is useful because it shows that the value function in our benchmark case is without loss of generality when mergers are frictionless. In particular, all our previous results apply:

Corollary 1. *Tournament bailouts credibly implement the first best as in Propositions 5 and 6 when mergers are frictionless. In particular, when $E > \underline{e}A$, frictionless mergers achieve the first best without bailouts.*

Tournaments implement the first best as the government is not forced to bail out the poorly performing banks, in spite of the value function (10). It can instead merge them with good banks, whose shareholders can be rewarded with additional funds if necessary.

4.2 Costly Mergers: Second Best Allocations

When mergers are costly ($\tau > 0$), in general the value function does not converge to the pure systemic risk model. Characterizing the second best allocation in that case is challenging, especially when we also consider endogenous bailouts. To make progress we specialize the value function (10) to

$$V = \bar{V} + v \sum_{i}^{N} a_i \min(0, y_i) \tag{12}$$

This is a conservative value function since it assumes no benefit from banks with capital surpluses $e_i > \underline{e}a_i$. In particular, the time inconsistency problem is extremely severe since, without mergers, it is ex post optimal to bail out *only* the banks with negative surplus.

The takeaway of this section is that even if mergers are costly, for low enough τ (below some threshold τ^*) the first best safety is credibly implementable, exactly as in the case of frictionless mergers. However, there is a stark discontinuity as the merger cost increases: τ above τ^* brings back full moral hazard. We start by analyzing the optimal ex post combination of mergers and bailouts before turning to the implications for ex ante safety incentives.

Ex Post Interventions: Mergers and Bailouts Define the surplus of good banks $Y_+ = \sum_{j \in N_+} a_j y_j$ and the shortfall of undercapitalized banks $Y_- = \sum_{j \in N_-} -a_j y_j$, hence $Y = Y_+ - Y_-$. With the value function (12) mergers are useful only if they tap into unused capital surplus. An immediate implication is that, if $\tau > 0$ and the value function is only weakly concave as in (12), then when $Y < 0$ the merger process will not lead to equalization of capital surpluses across banks. When $Y_+ > 0$ there is untapped capital surplus and the attractive merger target is the bank with the worst shortfall. Under

Assumption A2 we know that ${}^{d_i}/_{a_i} - r_i \geq 0$ hence $y_i \geq -\underline{e}$, so we make the following assumption to ensure that mergers are potentially useful:

Assumption. The merger cost satisfies $\tau < \underline{e}v$.

We can now describe the second best allocation. Define for $y \leq 0$ the cumulative shortfall function

$$\mathcal{Y}(y) = -\sum_{y_i \leq y} a_i y_i \in [0, Y_-]$$

and the following two cutoffs y_γ and y_τ:

$$y_\gamma = \inf_i y_i \text{ s.t. } \Gamma'(Y_- - \mathcal{Y}(y_i)) \leq v, \qquad \text{and} \qquad y_\tau = \sup_i y_i \text{ s.t. } \begin{cases} y_i \leq -\tau/v \\ \mathcal{Y}(y_i) \leq Y_+ \end{cases}$$

Both y_γ and y_τ are negative. To interpret y_γ, note first that is never ex post optimal to give more than $m_j = -a_j y_j$ to bank j. Thus if bailouts are the only option, the ex post efficient allocation that maximizes incentives gives a full bailout $m_j = -a_j y_j$ to banks starting from $y_j = 0$, until the marginal cost of further bailouts Γ' exceeds the marginal benefit v, which happens at y_γ. To interpret y_τ, note that if mergers are the only option, the ex post efficient allocation merges all the banks with y_i below y_τ to banks with a capital surplus, starting with the worst bank y_i. The two inequalities defining y_τ capture the fact that the merging process stops when either the marginal return to merging the next bank falls below the cost τ, or the entire capital surplus Y_+ has been exhausted. When $y_\tau > y_\gamma$, we define $y^* \in (y_\gamma, y_\tau)$ such that

$$y^* = \sup_i y_i \text{ s.t. } -y_i \Gamma'(Y_- - \mathcal{Y}(y_i)) \leq \tau.$$

Figure 1 illustrates the cumulative shortfall function $\mathcal{Y}$ before and after mergers, in the case $y_\tau < y_\gamma$.

When bailouts and mergers are both available, the ex post efficient allocation that maximizes incentives only bails out the least undercapitalized banks:

Lemma 4. *The efficient ex post policy combining bailouts and mergers is as follows:*

If $y_\tau \leq y_\gamma$ then (i) banks with $y_i \in [y_\gamma, 0]$ are fully bailed out ($m_i = -ay_i$) thus the aggregate bailout is $-\sum_{y_\gamma \leq y_i \leq y} a_i y_i = Y_- - \mathcal{Y}(y_\gamma)$; (ii) banks with $y_i \in (y_\tau, y_\gamma)$ are left untouched; (iii) banks with $y_i \leq y_\tau$ are merged with good banks.

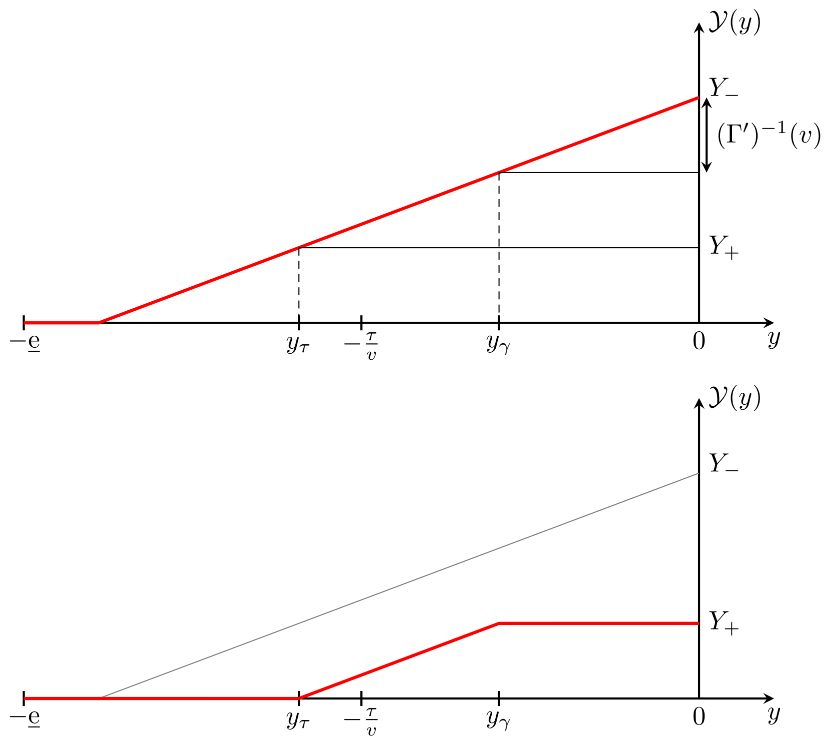

Figure 1: Cumulative shortfall function $\mathcal{Y}$ (in red) in the case $y_\tau < -\tau/v < y_\gamma$ before (top panel) and after (bottom panel) mergers and bailouts.

If $y_\tau > y_\gamma$ then (i) banks with $y_i \geq y^*$ are fully bailed out; (ii) banks with $y_i \leq y^*$ are merged with good banks.

Ex Ante Incentives We now show how mergers allow the government to reward good banks and punish weak ones and thus provide incentives, even when the loss function such as in (12) would call for fully bailing out weak banks absent the merger technology. Suppose the cost of funds is quadratic, $\Gamma(M) = \gamma \frac{M^2}{2}$. Then we can compute the thresholds as

$$y_\tau = -\frac{\tau}{v}, \quad y_\gamma = -\frac{v}{\gamma}, \quad y^* = -\sqrt{\frac{\tau}{\gamma}}.$$

Consider the case of 2 banks with the same shock structure as in Section 3.2, with a single crisis state s to simplify. There are four possible events, depending on which banks end up in groups H and L. If both banks end up with a capital surplus, no policy intervention is needed. If both banks end up with capital shortfalls ($y_1 = y_2 = s - \frac{d}{a} - \underline{e} < 0$) then there is no merger as $Y_+ = 0$. We assume the cost of funds is low enough that the only time-consistent policy in that case is to bailout both banks fully (so that both banks obtain a payoff $\underline{e}$), i.e., the following "joint bailout" condition holds:

$$v \geq 2\gamma \left(\frac{d}{a} + \underline{e} - s \right). \tag{13}$$

The most interesting case occurs if only bank 1, say, ends up in group H. Then in an equilibrium with safety choices x, $Y_+ = y_1 = s + x - \frac{d}{a} - \underline{e}$ and $Y_- = -y_2 = -\left(s - \frac{d}{a} - \underline{e}\right)$. Suppose that even under the laissez-faire safety $\tilde{x}$, the strong bank has enough capital to absorb the distressed one, that is, $\tilde{x} > 2\left(\frac{d}{a} + \underline{e} - s\right)$. Thus whenever bank 1 succeeds, a full merger is feasible. Condition (13) implies that if mergers are too costly, it is always optimal to fully bailout bank 2. However, if the cost of mergers τ is low enough

$$\tau \leq \tau^*(\gamma) = \gamma \left(\frac{d}{a} + \underline{e} - s \right)^2$$

then $y_2 \leq y^*$ and a merger is optimal. The merged bank's shareholders end up with 0, while the resulting equity of bank 1's shareholders is $2s + x^* - 2\frac{d}{a}$. Figure 2 illustrates how the thresholds y_τ, y_γ and y^* vary with the cost of mergers τ. A higher fiscal capacity (lower γ) makes bailouts more attractive and undermines the credibility of mergers, thus reducing τ^*.

Turning to ex ante incentives, we find a striking discontinuity in the merger cost τ:

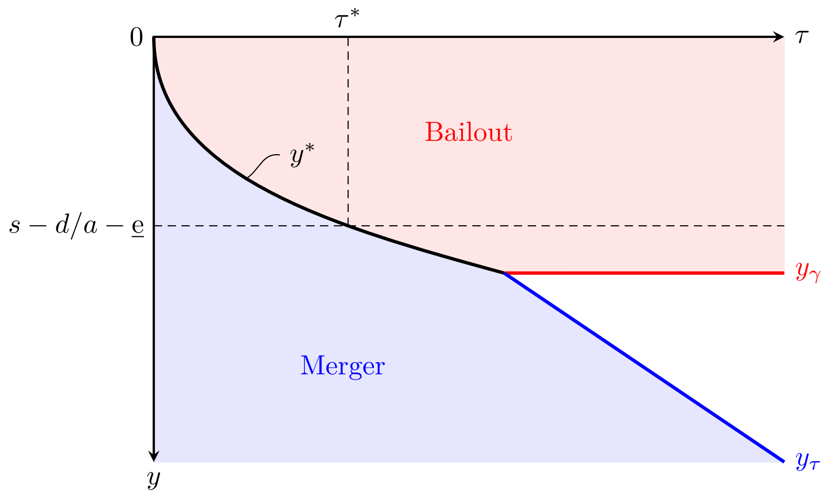

Figure 2: Optimal ex post policy as a function of the cost of mergers τ.

Proposition 10. *Suppose that γ is low enough.*[7] *For $\tau > \tau^*$ only the moral hazard safety level $\hat{x}$ (as defined in Proposition 3) is credibly implementable, while for $\tau \leq \tau^*$ the first best safety x^* is credibly implementable. Welfare decreases discontinuously at $\tau = \tau^*$.*

Proposition 10 extends our aggregation result under frictionless mergers to the case of costly mergers. We find that if τ is low enough then the first best is attainable, just like in the limit $\tau \to 0$. By contrast, for τ high enough we are back to the full moral hazard equilibrium. One interesting implication of this proposition is the complementarity between the efficiency of mergers at the micro level and the credibility of tournaments at the macro level. Efficient mergers allow the government to redistribute assets without creating aggregate costs, which creates scope to provide private incentives for safety even without commitment.

Finally, our proposal relies on rewarding strong banks, but we abstracted from imperfect information and the resulting fear of stigma that may prevent these strong banks from even accepting government support (Philippon and Skreta, 2012; Tirole, 2012). This was an important concern during the 2008 crisis: regulators had to force some of the healthier banks to accept government capital. First, we emphasize that the effective transfer must be attractive enough: banks reluctant to some forms of bailouts such as preferred stock injections will still welcome subsidized mergers with asset guarantees. Second, stigma is endogenous to the bailout mechanism, and our scheme works in the

[7] Formally $\gamma < \hat{\gamma}$ where $\hat{\gamma}$ is defined in the proof.

right direction. Under standard mechanisms allocating large bailouts to the weakest institutions, accepting public support is indeed a sign of weakness. But under a well-understood tournament mechanism, public support is a signal of strength instead. In fact, under our mechanism, upon the failure of some banks the market value of the surviving ones would rise, as they would now be revealed to be in the good camp.

5 Differentiated Banks: Too-Specific-To-Fail

In Section 4 we show how mergers can be used to optimally combine capital shortfalls and surpluses under the assumption that the same banking activities can be performed under different ownership structures. This may not be a good assumption when banks are geographically specialized and rely on soft information, or when the regulators worry about excessive local concentration in deposit taking as emphasized by Drechsler et al. (2014). Suppose then that banks are imperfectly substitutable and the value function is

$$V\left\{e_i + m_i\right\} = V\left(\phi\left\{e_i + m_i\right\} - \phi\left\{\underline{e}\right\}\right)$$

where

$$\phi\left\{e_i + m_i\right\} = \sum_{i=1}^{N} \left(e_i + m_i\right)^{\frac{\eta-1}{\eta}}$$

and $\eta > 1$ is the elasticity of substitution between banks. This value function converges to the one in the pure systemic model (5) as $\eta \to \infty$. It also captures the fact that it becomes more costly to take away the positive equity $e_i = a_i r_i - d_i$ from bank i as it gets smaller.

In this section we assume differentiability of f and V and a linear cost of funds $\Gamma\left(M\right) = \gamma M$ to simplify the exposition. Without commitment, perfect ex post efficiency requires equalizing the marginal return of transfers m_i across banks i, that is for each i

$$\frac{\eta - 1}{\eta} \left(e_i + m_i\right)^{\frac{-1}{\eta}} \frac{\partial V}{\partial e_i} \left\{e_i + m_i\right\} = \gamma.$$

Thus the government will fully insure all banks by setting the same level for ex post capital all banks $e_i + m_i = e_*$ irrespectively of individual bank performance, where $e_* > \underline{e}$ solves

$$\frac{\eta - 1}{\eta} e_*^{\frac{-1}{\eta}} V'\left\{e_*\right\} = \gamma \tag{14}$$

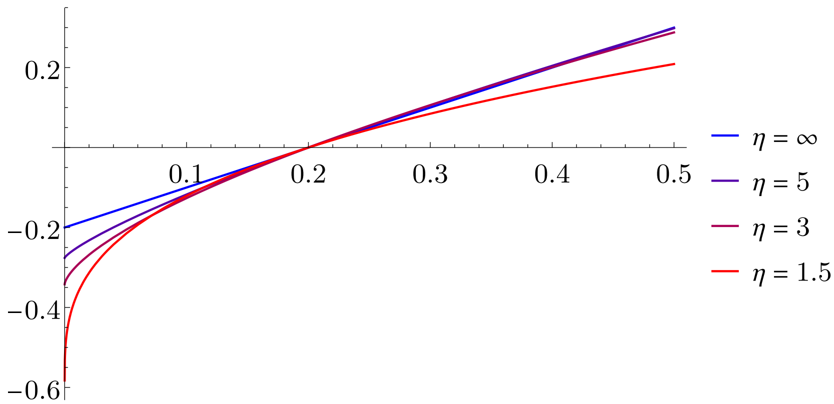

Figure 3: Function $V(e) = e^{\frac{\eta-1}{\eta}} - \underline{e}^{\frac{\eta-1}{\eta}}$ for different values of η. Lower η makes the function more concave, which increases incentives to offset individual capital shortfalls.

denoting $V'\{e_*\} = \frac{\partial V}{\partial e_i}\{e_*\}$.

At first glance, it seems that imperfect substitutability brings back the extreme form of moral hazard that arose under symmetric mechanisms. Each bank knows that it will be perfectly insured by the government since other banks will not be able to step in and replace it in case of resolution. In particular, our previous tournament scheme is not credible in this context. This extreme result comes from the extreme assumption that the government does not want to deviate at all from the ex post optimum. Indeed, if banks are almost perfectly substitutable ($\eta \to \infty$), imperfect insurance should have negligible costs and the model's conclusions should approach those of the pure systemic risk model.

We now relax the assumption of complete lack of commitment in two ways; both allow to re-establish our main result. In the first relaxation, we introduce a small amount of commitment, by giving the planner the ability to deviate slightly from the ex post optimum, by an amount at most $\epsilon > 0$ in welfare terms. We call this notion ϵ-commitment. In the second, and independent, relaxation, we consider a less stringent notion of time-consistency: the government can only deviate from promises to achieve a Pareto-improvement relative to the ex ante contract. The idea is that deviating from a pre-commitment is less likely to generate backlash and intense lobbying if all the involved parties benefit. In other words, we consider renegotiation-proof mechanisms in the language of Fudenberg and Tirole (1990). This solution concept provides a weak form of commitment consistent with the political economy of bailouts.

5.1 ϵ-Commitment

Consider a mechanism that transfers

$$m_i = e_* + d - r_i + \delta\left(r_i - \bar{r}\right) \tag{15}$$

to each bank so that the capital after bailout is $r_i - d + m_i = e_* + \delta\left(r_i - \bar{r}\right)$ where e_* is the ex post efficient (symmetric) capital that solves (14) and $\bar{r} = \frac{1}{N}\sum_i r_i$ is the average return. We are looking for a slope δ that is high enough to give incentives ex ante, while remaining low enough that the loss in ex post efficiency remains below some threshold ϵ. The next proposition shows how knife-edge the case of complete lack of commitment $\epsilon = 0$ is. In general, there is a trade-off between commitment and substitutability: with any small level of commitment $\epsilon > 0$, the first best is implementable if banks are sufficiently substitutable:

Proposition 11. *There exists $\alpha \in (0,1)$ increasing in ϵ such that the first best is implementable under ϵ-commitment using transfers*

$$m_i = e_* + d - r_i + \delta\left(r_i - \bar{r}\right)$$

with $\delta = \frac{1+\gamma}{1-\frac{1}{N}}$ if

$$\eta\epsilon \geq \frac{N}{\left(1-\frac{1}{N}\right)^2}\frac{\left(1+\gamma\right)^2\gamma\sigma_r^2}{2k\left(\gamma\right)\left(1-\alpha\right)}. \tag{16}$$

The right-hand side of (16) is increasing in γ, in the variance of returns σ_r^2, and in the number of banks N.

Equation (16) yields interesting comparative statics. The recurring theme in our paper is that once we allow for richer mechanisms, fiscal space (lower γ) is helpful for incentives. In this particular example, fiscal space and commitment ability ϵ are complement: fiscal space allows for larger bailouts and thus lower welfare losses from any ex post equity dispersion, as banks are dispersed around a level closer to the unconstrained optimum (that solves $V'\{e\} = 0$).

A contract with non-zero slope δ amplifies return differences arising from luck (in equilibrium), hence a lower variance of idiosyncratic risk σ_r^2 makes stronger incentives δ less costly to provide, which also decreases the amount of commitment needed. Finally, the number of banks N plays two roles: first, we impose the ϵ bound on the total welfare loss V, and a larger numbers of banks N increases any welfare loss mechanically: if ϵ-

efficiency applied to welfare *per bank* (i.e., $\Delta V \leq N\epsilon$) then $\bar{\delta}$ would be given by the same formula with $N = 1$; second, a larger N strengthens the incentive from δ. The first effect dominates.

Proposition 11 uncovers a novel policy implication for ex ante regulation. Existing policies, both micro- and macro-prudential, are focused on setting high enough capital and liquidity buffers, not so much on the scope of bank activities. But our model highlights the social cost of allowing banks to become "too-specific-to-fail". While the substitutability η must be taken as given ex post, there is a range of ex ante regulation that can effectively increase the substitutability η. For instance, even in settings where technological increasing returns to scale would call for having only one or two banks specialized in some activity (such as Bank of New York Mellon and JPMorgan Chase for the clearing of tri-party repos), credibility concerns give a rationale for imposing some redundancy. This insight is reminiscent of the industrial organization literature on multiple sourcing as a protection against ex post holdups (Shepard 1987, Farrell and Gallini 1988): a monopolist trying to encourage early product adoption may benefit from offering licenses to rivals, as a commitment to keep the post adoption market competitive.

5.2 Renegotiation-Proof Mechanisms

We now discuss another form of partial commitment. When banks are imperfect substitutes, their ex ante incentives are undermined by the lack of government commitment in two ways: ex post, the government would like to save the weakest banks, but it also doesn't want to favor the strong ones. Suppose, as in the literature on renegotiation-proof mechanisms, that it remains impossible to commit to ex post Pareto inefficient allocations, but that it it is politically costly to renege on promises when they end up hurting some subset of the agents. The interpretation is that banks (supported by their state or country if we interpret the imperfect substitutability as reflecting geographical segmentation) have a stronger incentive to lobby against an intervention if they have something to lose. As a result, the government will still help the worst banks (who have no reason to complain), but it is now able to credibly reward the strong banks.

To convey the point it is sufficient to consider the case of two banks $N = 2$. We assume that ex ante the government announces post recapitalization levels $(\bar{e}_1, \bar{e}_2)$ for the better and worse performing bank, respectively, such that ex post the government can choose its preferred allocation subject to the constraint that each bank must be weakly better off than under the contractual allocation $(\bar{e}_1, \bar{e}_2)$. Thus at date 1, given $(\bar{e}_1, \bar{e}_2)$ the government solves (suppose without loss that $r_1 > r_2$):

$$\max_{m_1, m_2} \quad V\left(\phi\left\{e_1 + m_i\right\} - \phi\left\{\underline{e}\right\}\right) - \gamma M$$
$$\text{s.t. } e_1 + m_1 \geq \bar{e}_1$$
$$e_2 + m_2 \geq \bar{e}_2$$

The following result shows that with enough fiscal capacity, the prospect of rewards is sufficiently strong to restore first best incentives, in the same spirit as our results on limited liability. To simplify, consider the additive return structure

$$r_i = x_i + s + \epsilon_i$$

and let $h = H'(0)$ where H is the c.d.f. of $\epsilon_2 - \epsilon_1$.

Proposition 12. *There exists $\hat{\gamma}$ such that for $\gamma < \hat{\gamma}$ the tournament contract $(\bar{e}_1, \bar{e}_2)$ where $\bar{e}_1$ is the unique solution to*

$$\frac{\partial \phi}{\partial e_2}\left(\bar{e}_1, \bar{e}_1 - \frac{1+\gamma}{h}\right) \times V'\left(\phi\left(\bar{e}_1, \bar{e}_1 - \frac{1+\gamma}{h}\right) - \phi\left(\underline{e}\right)\right) = \gamma \tag{17}$$

and $\bar{e}_2 = \bar{e}_1 - \frac{1+\gamma}{h}$ is renegotiation-proof and implements the first best safety x^*.

In the limit perfectly substitutable banks $\eta \to \infty$, the renegotiation-proof tournament converges to the tournament in Section 3. The renegotiation-proof "winner" payoff $\bar{e}_1$ (and therefore the payoff for the "loser" $\bar{e}_2 = \bar{e}_1 - \frac{1+\gamma}{h}$) increases as η decreases. The reason is that when banks are more specialized, it becomes less credible to punish the worst bank harshly. Ex post, the marginal benefit of bailing out the worst bank is higher when customers cannot easily switch to the best bank. Thus incentives must be provided through a better "carrot" for the better bank. Since the incentive condition pins down the payoff difference between the two banks, the worst bank also ends up with a larger bailout. The expected cost of ex post interventions $\mathbb{E}[m_1 + m_2] = 2\bar{e}_1 - \frac{1+\gamma}{h} - \mathbb{E}[r_1 + r_2]$ is thus higher when banks are more specialized.

Figure 4 shows a numerical example. As $\eta \to \infty$ the expected cost converges to the first best expected cost of bailouts (assuming banks all choose x^*) $\mathcal{K}(\gamma) - \mathbb{E}[r_1 + r_2]$. But note that the expected cost of intervention decreases quickly with η and becomes very close to the first best limit already when $\eta \approx 5$.

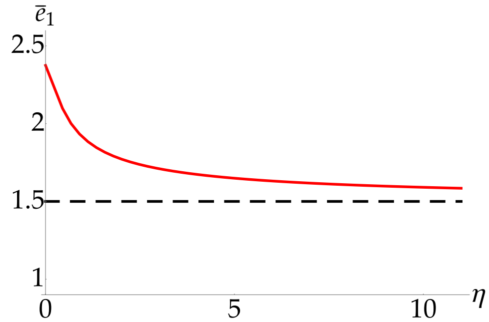

Figure 4: Renegotiation-proof prize $\bar{e}_1$ for the best bank as a function of the elasticity of substitution η. Dashed line: $\bar{e}_1$ with perfectly substitutable banks. Parameters: $V(x) = -\frac{x^2}{2}$, $\gamma = 0.5$.

6 Financial Contagion: Too-Interconnected-To-Fail

In this section we consider a different form of heterogeneity, arising from financial linkages between banks that generate comovement in returns. These linkages capture a variety of "contagion" forces, such as cross-exposures, fire sales, or domino effects, as studied in the financial networks literature (e.g., Caballero and Simsek 2013, Elliott et al. 2014, Acemoglu et al. 2015). The resulting return structure is significantly more complex than the one we have worked with so far: banks now have heterogeneous loadings on the aggregate risk factor s, and each bank is exposed to many other banks' idiosyncratic structural shocks ϵ_j.

We show how contagion leads to a natural notion of systemic risk: banks are more systemic when their performance has a stronger effect on the rest of the system. In turn, more systemic banks must act more prudently, and so a resolution mechanism must strive to give them stronger incentives. ex post, however, the government may consider these "super-spreader" banks too interconnected to fail (Haldane, 2013). Our main finding is that the constraints that financial linkages impose on bank resolution depend crucially on how bailout funds attributed to one bank spill over to other banks.

If public funds can be earmarked and bailout money cannot flow throughout the system to benefit other banks indirectly, our tournament mechanism remains credible and efficient under minor amendments. A bank's rank in the tournament is determined by its ex post performance, as in the baseline model, but now weighted by its systemic risk.

A subtle constraint appears if earmarking public funds is not possible, and bailout money can instead spillover to other banks. A first intuition would be that these spillover effects can reduce costs ex post, as it is now possible to rescue some banks indirectly, working through the linkages. The countervailing and dominating force, however, is that spillovers actually worsen the credibility problem. It becomes optimal to target the most systemic bank, as this is a cheap way to save the whole system. But this makes the moral hazard problem unsolvable, because the most systemic bank will now be completely insured and thus maximize risk-taking, thereby endangering the whole system.[8]

6.1 Restricted Bailouts

In this section, we focus on interconnectedness and simplify the other dimensions of the model, by assuming that all states s are systemic, and that f is differentiable. Suppose that conditional on a crisis, each bank i's return becomes a function of other banks j's returns through a linear relation:

$$\mathbf{r} = \mathbf{x} + \mathbf{s} + \boldsymbol{\epsilon} + \boldsymbol{\Omega}\mathbf{r}$$

with $\boldsymbol{\Omega} = \{\omega_{ij}\}$ where by convention $\omega_{ii} = 0$. We assume here that the interconnection between banks is based on *pre-bailout* returns r: at the ex post stage, bailouts do not spillover to other banks. The next subsection will consider the case in which bailout funds m_i cannot be "targeted" to bank i's shareholders, but also spill over to other banks j. As a result, returns can be solved as

$$\mathbf{r} = \boldsymbol{\Lambda}\left(\mathbf{x} + \mathbf{s} + \boldsymbol{\epsilon}\right) \tag{18}$$

where $\boldsymbol{\Lambda} = (\boldsymbol{I} - \boldsymbol{\Omega})^{-1}$. Call Λ_{ij} the elements of $\boldsymbol{\Lambda}$. The crisis value function in a contagion state becomes

$$V\left(\sum_i \lambda_i \left(x_i + s + \epsilon_i\right) + \sum_i m_i\right)$$

where $\lambda_i = \sum_j \Lambda_{ji}$ captures the *systemic risk* of bank i, that is how much other banks load on bank i's return, and thus how much bank i's return can affect the aggregate banking sector's shortfall through this form of financial contagion. Banks with higher

[8] In the knife-edge case in which multiple banks are equally systemic, we can still use a tournament within them and thus restore incentives.

weights λ_i are banks who have a high "network centrality": their returns have a relatively large impact on aggregate bank capital.

Suppose the cost of funds is linear hence the aggregate bailout is $\mathcal{M} = \mathcal{K}(\gamma) - R$. The results can readily be extended to a more general setting. The ex post optimality constraint remains unchanged: the total bailout has to satisfy $\sum_i m_i = \mathcal{M}$. The only difference in the first best allocation is that ex ante, more systemic banks should invest more in safety. For instance, when f is differentiable, the first best vector $\mathbf{x}^*$ solves

$$f'(x_i^*) = -\left(\frac{1-p_0}{p_0}\right)\lambda_i(1+\gamma). \tag{19}$$

Our baseline symmetric model is nested by setting $\mathbf{\Omega} = \mathbf{0}$ hence $\lambda_i = 1$ for all i. With heterogeneity, the first best requires that higher λ_i banks must invest in higher safety x_i^*.

While the most natural interpretation of contagion involves weights $\lambda_i > 1$ so that investment in safety by bank i has positive externalities on other banks' returns, note that nothing prevents weights λ_i from being lower than 1. This allows to capture in part negative actions that banks can take against their competitors, which become especially tempting in the presence of tournament incentives. In that case the first best solution is to reduce the investment x_i of such banks, and it can still be implemented through the handicapped tournament described below.

Handicapped Tournament. We show next that only slight modifications to our tournament mechanism are enough to accommodate the presence of this fairly general form our financial contagion. Intuitively, under heterogeneous systemic risk, the ex post bailout distribution must incentivize more systemic banks to hedge more. This is achieved by promising such banks higher prizes upon winning the tournament, or raising the effect of safety on their probability of "winning the tournament". An asymmetric or "handicapped" tournament contract can implement the first best, by simply ranking banks ex post according to their systemic-weighted performance $\tilde{\lambda}_i r_i$ instead of their raw return r_i. For simplicity, consider the case of two banks:

Proposition 13. *Suppose $N = 2$. Denote $h = H'(\lambda_1 x_1^* - \lambda_2 x_2^*)$ where H is the c.d.f. of $(\lambda_2 - \lambda_1)\eta + \lambda_2\epsilon_2 - \lambda_1\epsilon_1$, and*

$$\tilde{\lambda}_1 = \lambda_1 + \Lambda_{21} + \det\Lambda - 1, \qquad \tilde{\lambda}_2 = \lambda_2 + \Lambda_{12} + \det\Lambda - 1.$$

Then the following contract implements the first best (x_1^*, x_2^*) credibly:

If $\tilde{\lambda}_1 r_1 > \tilde{\lambda}_2 r_2$ then bank 1 obtains $m_1 = \frac{\mathcal{K}}{2} + \frac{1+\gamma}{2h} - r_1$ and bank 2 obtains $m_2 = \frac{\mathcal{K}}{2} - \frac{1+\gamma}{2h} - r_2$;

If $\tilde{\lambda}_2 r_2 > \tilde{\lambda}_1 r_1$ then bank 1 obtains $m_1 = \frac{\mathcal{K}}{2} - \frac{1+\gamma}{2h} - r_1$ and bank 2 obtains $m_2 = \frac{\mathcal{K}}{2} + \frac{1+\gamma}{2h} - r_2$.

Example. To illustrate the result, suppose bank 1 is systemic so $\omega_{21} = \omega \neq 0$ but bank 2 is not, $\omega_{12} = 0$. Then the matrix $\mathbf{\Lambda}$ is

$$\mathbf{\Lambda} = \begin{pmatrix} 1 & 0 \\ w & 1 \end{pmatrix}.$$

The weights λ_i that characterize the first best allocation through (19) are given by

$$\lambda_1 = 1 + w, \quad \lambda_2 = 1.$$

The weights $\tilde{\lambda}_i$ that make the first best allocation an equilibrium of the handicapped tournament are related but slightly different, given by

$$\tilde{\lambda}_1 = 1 + 2w, \quad \tilde{\lambda}_2 = 1.$$

If $\omega > 0$ as in the standard interpretation of contagion, the socially efficient allocation dictates that bank 1 invest more in safety in order to protect bank 2 indirectly. This higher safety can be induced through a tournament that makes it easier for bank 1 to earn the winning prize. If $\omega < 0$ instead, bank 1 has a negative externality on bank 2, and it is optimal to weaken its investment x_1 by under-weighting its performance in the tournament.

6.2 Contagious Bailouts

Finally, we consider the form of financial contagion that is hardest to overcome credibly. The regulator observes returns $\tilde{r}_i$ such that $\tilde{\mathbf{r}} = \Lambda(\mathbf{x} + \mathbf{s} + \boldsymbol{\epsilon})$ as in the previous subsection, before deciding on a bailout policy. The key difference is that now we suppose that bailout money itself is also "contagious". It is each bank j's post bailout equity $r_j + m_j$

(and not just r_j) that affects the value of other banks' assets r_i:

$$r_i = x_i + \sum_{j \neq i} \omega_{ij} \left(r_j + m_j \right) + s + \epsilon_i. \tag{20}$$

Adding m_i on each side and solving for $\mathbf{r} + \mathbf{m}$, we obtain in vector form

$$\mathbf{r} + \mathbf{m} = \Lambda \left(\mathbf{x} + \mathbf{m} + \mathbf{s} + \boldsymbol{\epsilon} \right) = \tilde{\mathbf{r}} + \Lambda \mathbf{m}.$$

The seemingly small difference relative to (18) turns out to be crucial in terms of policy implications. There is now an additional ex post asymmetry between banks: in the first best allocation, not only should more systemic banks (i.e., those with a higher λ_i) invest more in liquidity x ex ante; but as we will show, it is also efficient to focus the ex post government intervention on the most systemic bank. In the crisis state, the value function now writes

$$V \left(\sum_j \tilde{r}_j + \sum_i \lambda_i m_i \right)$$

The first best vector of safety $\mathbf{x}^*$ is the same as in the previous section. Ex post, however, since the shadow cost of public funds γ is the same for all banks i, a larger "bang for the buck" is obtained in terms of stabilizing the financial sector when the marginal dollar of public funds is allocated to the most systemic bank. Suppose that banks are strictly ranked according to their systemic risk, with bank 1 being the unique most systemic bank:

$$\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$$

and banks cannot be taxed to fund other banks, so that $m_i \geq 0$ (otherwise the result would be strengthened further, as the planner would then redistribute from banks $i \geq 2$ to bank 1). We have the following result regarding the optimal ex post intervention:

Lemma 5. *For any realization of pre-bailout returns $\tilde{\mathbf{r}}$, the optimal ex post policy is to transfer the full aggregate bailout $\mathcal{M}$ to bank 1: $m_1 = \mathcal{M}$, and nothing to other banks: $m_i = 0$ for all $i \geq 2$. The total bailout is $\mathcal{M} = \frac{\mathcal{K}}{\lambda_1} - \sum_{i=1}^{N} \tilde{r}_i$ and decreases with λ_1.*

For a given realization of returns, the loss is decreasing in the largest systemic weight λ_1. Ex post, it is cheaper to inject funds through the most systemic bank, and the more systemic it is, the cheaper the total cost of intervening. However, this will backfire ex ante: when bailouts are contagious, it becomes impossible to credibly punish bank 1 and reward other banks.

Proposition 14. *When bailout funds cannot be earmarked, the government has zero commitment, and banks are differentially interconnected, the equilibrium reverts to maximal risk-taking by the most systemic bank,* $x_1 = 0$*, and autarky-level risk-taking by other banks:* $x_i = \tilde{x}_i \quad \forall i \geq 2$*. The equilibrium bailout* $\mathcal{M} = \frac{\mathcal{K}}{\lambda_1} - \sum_{i=2}^{N} \lambda_i \tilde{x}_i - \sum_{i=1}^{N} \lambda_i (s + \epsilon_i)$ *exceeds the first best bailout by* $\mathcal{M} - M^* = \lambda_1 x_1^* + \sum_{i=2}^{N} \lambda_i (x_i^* - \tilde{x})$*, which is increasing with* λ_1*.*

The optimal aggregate bailout goes entirely to bank 1 and fully offsets bank 1's idiosyncratic shock ϵ_1; but it also depends on the realization of all the idiosyncratic shocks $\{\epsilon_j\}_{j>1}$. When some other banks do poorly, even if bank 1 has not suffered a negative idiosyncratic shock, the government still wants to inject more capital into the system. Banks hit by idiosyncratic shocks receive nothing because it is cheaper to inject the money through the most systemic bank.

The takeaway from this section is that financial contagion undermines credibility if and only if bailout funds can flow through the system and affect the performance of many banks besides the bank they are supposed to target. It is thus desirable to enforce a form of earmarking, where bailout money can be used to rescue specific institutions (in an asymmetric way, to provide incentives), but with some conditionality regarding its use. For instance, bailout funds should not be used primarily to repay debt to other banks (and it is always possible to bail out these downstream banks directly instead).

As a corollary, our model sheds new light on the "safe harbor" versus "automatic stay" debate.[9] It is of course well understood that safe harbor provisions can have negative effects on incentives for risk management (Roe, 2011; Bolton and Oehmke, 2014). Our model shows that the key issue is not the extent of moral hazard for the downstream banks, whose health is affected by systemic banks; it is instead that heterogeneity in systemic risk undermines commitment power, as it is not credible not to bail out the most systemic institutions even when they perform poorly. Once again, a key take-away from our analysis is the complementarity between micro regulations (such as the scope of

[9] Safe harbor provisions allow some creditors to walk away with their pledged collateral instead of joining the line of other creditors in the bankruptcy process. In bankruptcy creditors' claims on a failing firm are normally subject to "automatic stay". In this context, "safe harbor" is a super-seniority right that exempts some liabilities from automatic stay. Safe harbor rights were introduced in 1982 for repo contracts on treasuries but the Bankruptcy Abuse Prevention and Consumer Protection Act of 2005 added safe harbor provisions for repo loans based on mortgage collateral.

safe harbor provisions) and macro regulation (systemic risk management under limited commitment).

7 Conclusion

A standard takeaway of the literature is that without commitment, the government is powerless at providing incentives, hence moral hazard must ensue. Our paper goes against this common wisdom and proposes a way to bring back high-powered incentives, even in a world with no commitment, by using tournaments.

Of course, once we do that, we also bring back the potential pitfalls of high-powered incentives, as in the multitask framework of Holmström and Milgrom (1991). Tournaments may induce banks to manipulate the return measures serving as inputs in the mechanism, or to take actions undermining other banks' performance. Yet if such issues arise, they would signal the success of our scheme at overcoming the basic moral hazard problem, and could be corrected by dampening incentives. Indeed, we considered such an example in the context of financial contagion, showing how to properly handicap the tournament when a bank imposes a negative externality on the system.

For theoretical clarity, we mostly framed the implementation in terms of standard taxes and transfers. A broader interpretation is that the government should lean towards policies that reward strong performers and punish weak ones, e.g., through differentiated restrictions on executive compensation, clawbacks as we discussed in the context of limited liability, or delays in bailouts and access to credit facilities.

References

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2015), "Systemic Risk and Stability in Financial Networks", *American Economic Review*, Vol. 105, pp. 564–608.

Acharya, Viral, Itamar Drechsler, and Philipp Schnabl (2014), "A Pyrrhic Victory? - Bank Bailouts and Sovereign Credit Risk", *Journal of Finance*, Vol. 69, pp. 2689–2739.

Acharya, Viral V, Lasse Heje Pedersen, Thomas Philippon, and Matthew Richardson (2016), "Measuring Systemic Risk", *The Review of Financial Studies*.

Acharya, Viral V. and Tanju Yorulmazer (2007), "Too many to fail: An analysis of time-inconsistency in bank closure policies", *Journal of Financial Intermediation*, Vol. 16, pp. 1–31.

——— (2008), "Cash-in-the-Market Pricing and Optimal Resolution of Bank Failures", *The Review of Financial Studies*, Vol. 21, pp. 2705–2742.

Bolton, Patrick and Martin Oehmke (2014), "Should Derivatives be Privileged in Bankruptcy?", *Journal of Finance*.

Caballero, Ricardo J. and Alp Simsek (2013), "Fire Sales in a Model of Complexity", *The Journal of Finance*, Vol. 68, pp. 2549–2587.

Chari, V. V. and Patrick J. Kehoe (2016), "Bailouts, Time Inconsistency and Optimal Regulation: A Macroeconomic View", *American Economic Review*, Vol. 106, pp. 2458–2493.

Clayton, Christopher and Andreas Schaab (2021), "Bail-Ins, Optimal Regulation, and Crisis Resolution", working paper.

Dam, Lammertjan and Michael Koetter (2012), "Bank Bailouts and Moral Hazard: Evidence from Germany", *The Review of Financial Studies*, Vol. 25, pp. 2343–2380.

Dávila, Eduardo and Ansgar Walther (2020), "Does size matter? Bailouts with large and small banks", *Journal of Financial Economics*, Vol. 136, pp. 1–22.

Demange, Gabrielle (2020), "Resolution rules in a system of financially linked firms", March, working paper or preprint.

Dewatripont, Mathias and Jean Tirole (2018), "Liquidity Regulation, Bail-ins and Bailouts",Technical report.

Diamond, Douglas W. and Philip H. Dybvig (1983), "Bank Runs, Deposit Insurance, and Liquidity", *Journal of Political Economy*, Vol. 91, pp. 401–419.

Diamond, Douglas W. and Raghuram G. Rajan (2011), "Fear of Fire Sales, Illiquidity Seeking, and Credit Freezes", *Quarterly Journal of Economics*, Vol. 126, pp. 557–591.

——— (2012), "Illiquid Banks, Financial Stability, and Interest Rate Policy", *Journal of Political Economy*, Vol. 120, pp. pp. 552–591.

Drechsler, Itamar, Philipp Schnabl, and Alexi Savov (2014), "The Deposits Channel of Monetary Policy The Deposits Channel of Monetary Policy", NYU.

Elliott, Matthew, Benjamin Golub, and Matthew O. Jackson (2014), "Financial Networks and Contagion", *American Economic Review*, Vol. 104, pp. 3115–53.

Farhi, Emmanuel and Jean Tirole (2012), "Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts", *American Economic Review*, Vol. 102, pp. 60–93.

Farrell, Joseph and Nancy Gallini (1988), "Second-Sourcing as a Commitment: Monopoly Incentives to Attract Competition", *The Quarterly Journal of Economics*, Vol. 103, pp. 673–694.

Financial Stability Board (2021), "Evaluation of the effects of too-big-to-fail reforms".

Fudenberg, Drew and Jean Tirole (1990), "Moral Hazard and Renegotiation in Agency Contracts", *Econometrica*, Vol. 58, pp. 1279–1319.

Haldane, Andrew G (2013), *Rethinking the financial network*, pp. 243–278, Wiesbaden: Springer Fachmedien Wiesbaden.

Hart, Oliver and John Moore (1995), "Debt and Seniority: An Analysis of the Role of Hard Claims in Constraining Management", *American Economic Review*, Vol. 85, pp. 567–585.

Holmström, Bengt and Paul Milgrom (1991), "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design", *JL Econ. & Org.*, Vol. 7, p. 24.

Holmström, Bengt and Jean Tirole (1997), "Financial Intermediation, Loanable Funds and the Real Sector", *Quarterly Journal of Economics*, Vol. 112, pp. 663–691.

Innes, Robert (1990), "Limited Liability and incentive contracting with ex ante action choices", *Journal of Economic Theory*, Vol. 52, pp. 45–67.

Kasa, Kenneth and Mark M. Spiegel (2008), "The Role of Relative Performance in Bank Closure Decisions", *FRBSF Economic Review*.

Keister, Todd (2016), "Bailouts and Financial Fragility", *Review of Economic Studies*, Vol. 83, pp. 704–736.

Keister, Todd and Yuliyan Mitkov (2021), "Allocating Losses: Bail-ins, Bailouts and Bank Regulation", January.

Kelly, Bryan, Hanno Lustig, and Stijn Van Nieuwerburgh (2016), "Too-Systemic-To-Fail: What Option Markets Imply About Sector-wide Government Guarantees", *American Economic Review*.

Lazear, Edward P. and Sherwin Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contracts", *Journal of Political Economy*, Vol. 89, pp. 841–864.

Malherbe, Frederic (2014), "Self-Fulfilling Liquidity Dry-Ups", *The Journal of Finance*, Vol. 69, pp. 947–970.

Milgrom, Paul and Christina Shannon (1994), "Monotone Comparative Statics", *Econometrica*, Vol. 62, pp. 157–180.

Myers, Stewart C (1977), "Determinants of Corporate Borrowing", *Journal of Financial Economics*, Vol. 5, pp. 147–175.

Nachman, David and Thomas Noe (1994), "Optimal Design of Securities under Asymmetric Information", *Review of Financial Studies*, Vol. 7, pp. 1–44.

Nosal, Jaromir B. and Guillermo Ordonez (2016), "Uncertainty as commitment", *Journal of Monetary Economics*, Vol. 80, pp. 124–140.

Perotti, Enrico C and Javier Suarez (2002), "Last bank standing: What do I gain if you fail?", *European Economic Review*, Vol. 46, pp. 1599–1622, Symposium on Finance.

Philippon, Thomas and Philipp Schnabl (2013), "Efficient Recapitalization", *Journal of Finance*, Vol. 68, pp. 1–42.

Philippon, Thomas and Vasiliki Skreta (2012), "Optimal Interventions in Markets with Adverse Selection", *American Economic Review.*

Roe, Mark J (2011), "The Derivative Market's Payments Priorities as Financial Crisis Accelerator", *Stanford Law Review*, Vol. 63, pp. 539–590.

Sapienza, Paola (2002), "The Effects of Banking Mergers on Loan Contracts", *Journal of Finance*, Vol. 57, pp. 329–367.

Shepard, Andrea (1987), "Licensing to Enhance Demand for New Technologies", *RAND Journal of Economics*, Vol. 18, pp. 360–368.

Stein, Jeremy (2012), "Monetary Policy as Financial-Stability Regulation", *The Quarterly Journal of Economics*, Vol. 127, pp. 57–95.

Stein, Jeremy C (2002), "Information Production and Capital Allocation: Decentralized versus Hierarchical Firms", *Journal of Finance*, Vol. 57, pp. 1891–1921.

Tirole, Jean (2006), *The Theory of Corporate Finance*: Princeton University Press.

——— (2012), "Overcoming Adverse Selection: How Public Intervention Can Restore Market Functioning", *American Economic Review*, Vol. 102, pp. 29–59.

Topkis, Donald M. (1978), "Minimizing a Submodular Function on a Lattice", *Operations Research*, Vol. 26, pp. 305–321.

Appendix

A Micro-foundations for V and $\underline{e}$

Our model's value function V is meant to capture, in a tractable and unified way, a variety of externalities that arise when banks are solvent but poorly capitalized. The general formulation also highlights throughout the paper which key features matter for the provision of incentives, e.g., the degree of differentiation between banks. Nevertheless, in this section we give two (non-exclusive) illustrations. The first example focuses on banks' liability side, through the money market disturbances that happen when haircuts are imposed on creditors. The second example focuses on banks' asset side: new investment opportunities can emerge even during a crisis, but limited pledgeability prevents banks from realizing these investments unless they bring enough equity/liquidity into these states.

Money market instability. Suppose that when a bank's equity falls below a threshold $\underline{e}a_i$, creditors start running, unless the equity is replenished to $\underline{e}a_i$. The costs of allowing for a run are too high (e.g., the illiquidity discount on assets in place is too large), so banks must find a way to reach $\underline{e}a_i$. In the short run it is difficult to do it by issuing new shares, hence absent bailouts the only way to raise equity is to renegotiate the existing debt down, to a new level $\tilde{d}_i$ such that $a_i r_i - \tilde{d}_i = \underline{e}a_i$ that is

$$\tilde{d}_i = a_i r_i - a_i \underline{e}.$$

The renegotiation is approximately costless from the bank's private viewpoint, so that banks do not self-insure against these run events and only care about returns. But renegotiation is socially costly, as it creates a financial stability externality

$$\phi\left(d_i - \tilde{d}_i\right) = \phi\left(\underline{e}a_i - e_i\right)$$

where ϕ is increasing and weakly convex. For instance, if money market funds are highly exposed to banks' commercial paper, a debt write-down may trigger a run on money market funds and further instability in money markets. The cost ϕ indexed how "bailinable" the debt d_i is. Note that our goal here is not to provide deep foundations for limited bailinability: in practice this is a constraint taken as given by regulators, and related to holdout problems or incomplete contracts. Summing over all banks, the resulting value function is

$$V = -\sum_i \phi\left(\underline{e} a_i - e_i\right).$$

Whether ϕ is concave or linear, and thus how good an approximation the pure systemic risk provides, depends on other features of money markets, such as how diversified the money market funds are. ϕ will be more concave if some funds' holdings are extremely concentrated in some particular banks' debt, such as when the Reserve Primary Fund broke the buck due to its exposure to Lehman's commercial paper in 2008. ϕ will be closer to linear if funds are well-diversified, as then the aggregate debt write-down will be the most relevant variable.

New bank investments and limited pledgeability. Another natural foundation comes from a standard model with liquidity shocks and limited pledgeability à la Holmstrom Tirole. Banks have new investment opportunities (or equivalently liquidity shocks they need to cover), which they can finance by borrowing against their future equity. If equity is too low, even solvent banks will be constrained in their reinvestment scale, which generates an externality V if the social planner cares about these projects.

Concretely, we unfold our baseline model's date $t = 1$ into an intermediate date $t = 1$ and a final date $t = 2$. At the beginning of $t = 1$, banks' assets in place a_i that mature at $t = 2$ have a value $a_i r_i$ while debt d_i is also due at $t = 2$, so the value of their equity at the beginning is $e_i = a_i r_i - d_i$. There is a large supply of new investment opportunities: an investment k_i at $t = 1$ produces output $f(k_i)$ at $t = 2$ where f is weakly concave.

Banks must issue new debt l_i at some competitive rate ρ to finance these new investments. There is an upward sloping aggregate debt supply curve $L(\rho)$. Assume the output from these new investments is not pledgeable at all, while the output from the assets in place is fully pledgeable. For instance, if limited pledgeability arises from a model of moral hazard and private benefits, the assets in place may not require monitoring or screening effort anymore once at $t = 1$, unlike the new investments. More generally, as long as the proceeds from the assets in place are somewhat pledgeable and the new projects are not perfectly pledgeable, equity e_i may play a role to relax the date-1 financial constraint (Tirole, 2006). Banks solve

$$\begin{aligned} &\max f(k_i) - \rho l_i \\ &\text{s.t. } k_i \le e_i + m_i \\ &\quad\;\; k_i = l_i + m_i \end{aligned}$$

For a given rate ρ the unconstrained level of investment $\bar{k}$ solves

$$f'\left(\bar{k}(\rho)\right) = \rho$$

$\bar{k}(\rho)$ is decreasing in ρ if f is strictly concave; if f is linear equal to $f(k) = \rho_1 k$ then $\bar{k} = k_{max}$ if $\rho < \rho_1$ and can take any positive value if $\rho = \rho_1$.

Given the credit constraint the investment of bank i is thus

$$k_i = \min\left\{e_i + m_i, \bar{k}\right\}.$$

If the social planner values the return on new projects k_i we can express the value function V as

$$V\{e_i + m_i\} = \sum_i \min\left\{f\left(\bar{k}(\rho)\right), f(e_i + m_i)\right\}$$

where ρ itself depends on the vector $\{e_i + m_i\}$ and is determined by the market clearing condition for bank debt issued at $t = 1$:

$$L(\rho) = \sum_i \left(\min\left\{\bar{k}(\rho), e_i + m_i\right\} - m_i\right).$$

The simpler case of an exogenous interest rate ρ^* is nested, corresponding to a perfectly elastic supply curve $\rho = \rho^*$.[10] When f is linear (more generally, when decreasing returns are not at the bank level but at the aggregate level through $f\left(\sum k_i\right)$) the value function simplifies to

$$V = \min\left\{L(\rho_1), \sum_i (m_i + e_i)\right\}.$$

The maximal possible aggregate reinvestment is attained when all N banks are uncon-

[10] For general L, one can show that even taking into account the general equilibrium feedback on ρ, V remains increasing in e_i and it is concave if f is concave enough.

strained. It is given by $\bar{K} = L(\bar{\rho})$ where the maximal interest rate $\bar{\rho}$ solves

$$\bar{\rho} = f'\left(\frac{L(\bar{\rho})}{N}\right)$$

When f is linear then $\bar{\rho} = \rho_1$. Thus as in our baseline model, there is a level $\underline{\mathrm{e}} = \frac{L(\bar{\rho})}{N}$ such that there is no externality (V does not increase with e_i) if all banks have equity $e_i \geq \underline{\mathrm{e}}$.

B Tournaments with Heterogeneous Bank Size

In the general case with different bank sizes a_i, bank i chooses its safety investment to solve:

$$\hat{x}_i = \arg\max_{x_i \geq 0} p_0 f\left(x_i\right) + \left(1 - p_0\right)\left(\mathbb{E}\left[r_{i,s} \mid x_i\right] + \mathbb{E}\left[\frac{m_{i,s}\left(\mathbf{r}\right)}{a_i} \mid \mathbf{x}\right]\right). \tag{21}$$

Suppose the conditions of Lemma 2 hold hence the first best safety x^* does not depend on size. Importantly, due to the credibility constraint the reward Δ in the bonus-malus tournament cannot depend on size either: the gain of one bank is the loss of another. But if the tournament rule only compares raw returns to determine who wins and who loses, larger banks will in general choose a lower level of safety than smaller banks, because the potential prize Δ is smaller as a fraction of their assets.

We can solve this issue by considering the following handicapped tournament

$$m_i = \begin{cases} \frac{a_i}{A}\mathcal{M}\left(K - R\right) + \Delta & \lambda_i r_{i,s} > \lambda_j r_{j,s} \\ \frac{a_i}{A}\mathcal{M}\left(K - R\right) - \Delta & \lambda_i r_{i,s} < \lambda_j r_{j,s} \end{cases} \tag{22}$$

that compares weighted returns $\lambda_i r_i$ instead of raw returns to determine the bailout allocation. Given $\lambda = \frac{\lambda_1}{\lambda_2}$ the best response function for bank 1 is

$$\hat{x}_1 = \beta_1\left(\Delta, \lambda, x_2\right) = \arg\max_{x_1} p_0 f\left(x_1\right) + \left(1 - p_0\right)\left(\mathbb{E}\left[r_{1,s} \mid x_1\right] + \Omega\left(x_1, x_2\right)\right) + 2\frac{\Delta}{a_1}\int_s \mathbb{P}\left[\lambda r_{1,s} > r_{2,s} | \mathbf{x}\right] p_s ds,$$

while the best response function for bank 2 is

$$\hat{x}_2 = \beta_2\left(\Delta, \lambda, x_1\right) = \arg\max_{x_2} p_0 f\left(x_2\right) + \left(1 - p_0\right)\left(\mathbb{E}\left[r_{2,s} \mid x_2\right] + \Omega\left(x_1, x_2\right)\right) - 2\frac{\Delta}{a_2}\int_s \mathbb{P}\left[\lambda r_{1,s} > r_{2,s} | \mathbf{x}\right] p_s ds.$$

We thus look for a pair Δ, λ that implements the first best:

$$x^* = \beta_1\left(\Delta, \lambda, x^*\right)$$
$$x^* = \beta_2\left(\Delta, \lambda, x^*\right)$$

To characterize when this is possible, we use a more specific example of returns:

$$r_i = x_i + s + \epsilon_i. \tag{23}$$

Then

$$\mathbb{P}\left[\lambda x_1 - x_2 > (1-\lambda)\, s + \epsilon_2 - \lambda \epsilon_1\right] = H_s\left(\lambda x_1 - x_2; \lambda\right)$$

where $H_s(\cdot;\lambda)$ is the c.d.f. of $(1-\lambda)\, s + \epsilon_2 - \lambda\epsilon_1$. The marginal incentives from the tournament for banks 1 and 2 are respectively

$$\frac{\partial}{\partial x_1}\left(2\frac{\Delta}{a_1}\int_s H_s\left(x_1, x_2; \lambda\right) p_s ds\right) = 2\Delta\frac{\lambda}{a_1}\int_s H_s'\left(\lambda x_1 - x_2; \lambda\right) p_s ds$$

$$\frac{\partial}{\partial x_2}\left(-2\frac{\Delta}{a_2}\int_s H_s\left(x_1, x_2; \lambda\right) p_s ds\right) = 2\frac{\Delta}{a_2}\int_s H_s'\left(\lambda x_1 - x_2; \lambda\right) p_s ds.$$

so as long as $\int_s H_s'\left(\lambda x_1 - x_2; \lambda\right) p_s ds > 0$ there exists a λ such that the two banks to choose the same x^*.

Note that the condition $\int_s H_s'\left(\lambda x_1 - x_2; \lambda\right) p_s ds > 0$ imposes an upper bound on the relative size of the two banks. If a_1/a_2 is too large, then no λ can generate first best incentives for the larger bank and we are back to the moral hazard unavoidable in a one-bank world.

Proposition 15. *Suppose that $N = 2$, $a_1 \geq a_2$, and returns follow (23) with ϵ_i distributed over $[0, \bar{\epsilon}]$. Then there exists*

$$\kappa \in \left(0, \frac{\bar{\epsilon}}{x^* + \inf s}\right)$$

such that a handicapped tournament (22) can implement the first best safety if and only if

$$\frac{a_1}{a_2} < 1 + \kappa.$$

Proof. $\lambda = 1$ implements the first best as $\frac{a_1}{a_2} \to 1$. For $\lambda = \frac{a_1}{a_2}$ the tournament incentives are the same while $\frac{\partial \Omega}{\partial x_1} < \frac{\partial \Omega}{\partial x_2}$ hence bank 1 chooses a lower safety than bank 2. Hence we need $\lambda > \frac{a_1}{a_2}$. We can compute

$$H_s(\lambda x_1 - x_2; \lambda) = \int_0^{\bar{\epsilon}} G_\epsilon(\lambda\epsilon_1 + \lambda x_1 - x_2 - (1-\lambda)s)\, d\epsilon_1$$
$$H_s'(\lambda x_1 - x_2; \lambda) = \int_0^{\bar{\epsilon}} g_\epsilon(\lambda\epsilon_1 + \lambda x_1 - x_2 - (1-\lambda)s)\, d\epsilon_1$$

where G_ϵ and g_ϵ are the c.d.f. and p.d.f. of ϵ_1, respectively. Then for $x_1 = x_2 = x^*$

$$\lambda\epsilon_1 + \lambda x_1 - x_2 - (1-\lambda)s \leq \bar{\epsilon} \Leftrightarrow \epsilon_1 \leq \frac{\bar{\epsilon} - (\lambda - 1)(x^* + s)}{\lambda}$$

Therefore

$$\int_s H_s'(\lambda x_1 - x_2; \lambda)\, p_s ds = \int_s \left(\int_0^{\bar{\epsilon}} g_\epsilon(\lambda\epsilon_1 + \lambda x_1 - x_2 - (1-\lambda)s)\, d\epsilon_1 \right) p_s ds$$

is negative if $\lambda > 1 + \frac{\bar{\epsilon}}{x^* + \inf s}$. This shows that if $\frac{a_1}{a_2} > 1 + \frac{\bar{\epsilon}}{x^* + \inf s}$ the handicapped tournament cannot implement the first best. □

C Proofs

Proof of Proposition 1. First note that if $R > K$ the solution is obviously $M = 0$. We can therefore restrict our attention to $R < K$ and $M \geq 0$. Because V is concave. The solution $x^*(\theta, \kappa)$ to the problem $\max_x f(x - \theta) + g(k - x)$ where f and g are concave is increasing in θ and κ with slopes less than one, i.e., such that $x^* - \theta$ is decreasing in θ and $k - x^*$ is increasing in k. Therefore $\mathcal{M}(R, K)$ is increasing in $K - R$ with slope less than one. The comparative statics with respect to γ come directly from the fact that $\Gamma(M; \gamma)$ is increasing and super-modular. The fact that $\mathcal{V}$ is concave comes from the fact that V is concave and the fact that $\mathcal{M}$ has a slope less than 1. The definition of $\mathcal{K}(\gamma)$ is the same as in the next example.

Proof of Lemma 1. First note that if $R > K$ the solution is obviously $M = 0$. We can therefore restrict our attention to $R < K$ and $M \geq 0$. To exploit the quasi-linear preferences we change variable from M to $\hat{M} \equiv M + R - K$. We can rewrite the loss minimization problem (6) as

$$\max_{\hat{M} \geq R-K} V\left(\hat{M}\right) - \gamma\left(\hat{M} + K - R\right)$$

If $\hat{M} = R - K$ the solution is $M = 0$. If $\hat{M} > R - K$, then it solves

$$\hat{M}(\gamma) = \arg\max_{\hat{M}} \left\{ V\left(\hat{M}\right) - \gamma \hat{M} \right\}$$

which is negative and decreasing in γ. Since $M = \hat{M} + K - R$, we then get $M = \mathcal{K}(\gamma) - R$ with $\mathcal{K}(\gamma) = \hat{M}(\gamma) + K$. Putting the two cases together, we therefore get $M = \max\{0, \mathcal{K}(\gamma) - R\}$.

Proof of Lemma 2. Suppose G_ϵ does not depend on x. Define $\bar{r}(x, s) = \mathbb{E}[r_{i,s} \mid x, s]$. We have

$$\begin{aligned}
\mathbf{x}^* = \arg\max_{\mathbf{x} \geq 0} p_0 \sum_i f(x_i) + (1 - p_0) \int_s \sum_i \bar{r}(x_i, s)\, dP(s) \\
+ \frac{1}{a_i} \int_s dP(s) \int_\varepsilon \mathcal{V}\left(\sum_i a_i \bar{r}(x_i, s) + \varepsilon - K\right) d\bar{G}_\epsilon(\varepsilon)
\end{aligned}$$

where $\bar{G}_\epsilon(\varepsilon)$ is the convolution of the distributions G_ϵ. It does not depend on $\mathbf{x}$. Therefore

$$\frac{1}{a_i}\frac{\partial}{\partial x_i}\mathbb{E}\left[\mathcal{V}(R)\mid \mathbf{x}, s\right] = \bar{r}_x(x_i, s)\,\mathbb{E}\left[\mathcal{V}'(R)\mid \mathbf{x}, s\right]$$

and the optimal choice of x_i does not depend on the size of bank i.

Proof of Lemma 3. We use the standard notations $R_{-i} = \sum_{j\neq i} a_j r_{j,s}$ and

$$\begin{aligned}
\Phi_N(R\mid \mathbf{x}) &= \mathbb{P}\left(\tilde{R} < R \mid \mathbf{x}\right) \\
&= \int_s \mathbb{P}\left(\sum_{i=1}^{N} a_i r_{i,s} < R \mid \mathbf{x}, s\right) p_s ds \\
&= \int_s \mathbb{P}\left(a_1 r_{1,s} < R - R_{-1} \mid \mathbf{x}, s\right) p_s ds \\
&= \int_s \int_{R_{-1}} G\left(\frac{R - R_{-1}}{a_1} \mid x_1, s\right) d\Phi_{N-1}\left(R_{-1} \mid \mathbf{x}_{-1}, s\right) p_s ds
\end{aligned}$$

Since $G(. \mid x_i, s)$, is decreasing in x_i, so is $\Phi_N(R\mid \mathbf{x})$. Since $\mathcal{M}$ is decreasing in R, $\Omega(x_i; \mathbf{x}_{-i})$ in decreasing in x_i for any i. Since $G(. \mid x, s)$ is $\mathcal{C}^1$ in x we have

$$\frac{\partial \Phi_N(R\mid \mathbf{x})}{\partial x_i} = \int_s \int_{R_{-1}} \frac{\partial G\left(\frac{R-R_{-1}}{a_1} \mid x_i, s\right)}{\partial x_i} d\Phi_{N-1}\left(R_{-i} \mid \mathbf{x}_{-i}, s\right) p_s ds$$

is negative and increasing in $\mathbf{x}_{-i}$ since $\Phi_{N-1}(. \mid \mathbf{x}_{-i}, s)$ is decreasing in $\mathbf{x}_{-i}$. Therefore $\frac{\partial \Omega}{\partial x_i}$ is increasing in $\mathbf{x}_{-i}$.

Proof of Proposition 3. (i) Because $\frac{\partial \Omega}{\partial x_i}$ is increasing in x_{-i}. (ii) Because Ω is decreasing. (iii) Because $\mathcal{M}$ is decreasing in γ hence Ω is super-modular in (x_i, γ). (iv) follows from the fact that f is maximized at $x = 0$.

Proof of Proposition 4. The objective function is super-modular in (x_1, Δ) since H is increasing in x_1 therefore x_1 is increasing in Δ. Suppose that $x_2 = x^*$. Clearly $\hat{x}_1(0, x^*) < x^*$. On the other $\lim_{\Delta\to\infty} x_1(\Delta, x^*) = 1$. Since x_1 is continuous there is a unique Δ^* such that $x_1(\Delta^*, x^*) = x^*$. The same holds for x_2 by symmetry.

Proof of Proposition 6. The proof has two steps. Let x^{max} be the maximum implementable level of safety. The first step is that the planner can always improve upon purely private incentives. Any bailout function with $m_i(r_i < \text{med}(\mathbf{r})) = 0$ and $m_i(r_i > \text{med}(\mathbf{r})) = 2\mathcal{M}/N$ satisfies $\hat{x}_i > \tilde{x}$. Therefore $x^{max} > \tilde{x}$. The second step is that when $\gamma \to 0$ the government can fully insure downside risk: $\lim_{\gamma\to 0}\mathcal{V} = \bar{V}$ and $\lim_{\gamma\to 0} x^* = \tilde{x}$. Therefore $\lim_{\gamma\to 0} x^* < x^{max}$.

Proof of Proposition 7. Let us consider the implementation of a symmetric equilibrium x. When all banks make the same choice the aggregate return does not depend on the random selection of the groups H and L:

$$R = A(s + hx).$$

where $h = N_H/N$ is the probability that any particular bank ends up in group H. In particular, the first best solves

$$x^* = \arg\max_{x\geq 0} p_0 Af(x) + (1-p_0)A(s+hx) + \mathbb{E}[\mathcal{V}(As + Ahx - K)].$$

Let us now consider incentive constraints. If one bank deviates, the aggregate return depends both on s and on the group selection. Define $\tilde{X}_H = \sum_{i\in H} x_i$. The return is then

$$\tilde{R} = As + ah\tilde{X}_H.$$

If a bank deviates it will choose $x = 0$ so that it can hide among the legitimate banks of group L. Because banks' outcomes are binary, the bailout takes the simple form $m_L = a\mu_L$ for group $\tilde{L}$ (banks with low returns) and $m_H = a\mu_H$ to group $\tilde{H}$ (banks with high returns). Therefore the credibility constraint is

$$\left(N - \tilde{N}_H\right) a\mu_L + \tilde{N}_H a\mu_H = \mathcal{M}\left(K - \tilde{R}\right).$$

Consider the incentive constraint of bank 1 given that all the other banks play x^*. If bank 1 chooses $x_1 = x^*$ its expected payoffs are

$$p_0 af(x) + (1-p_0)a(\bar{s} + hx + \mathbb{E}[(1-h)\mu_L(s) + h\mu_H(s)])$$

If bank 1 instead chooses $x_1 = 0$ its expected payoffs are

$$p_0 af(0) + (1-p_0)a(\bar{s} + \mathbb{E}[(1-h)\mu_L(s) + h\mu_L(s, N_H - 1)])$$

because ex post with probability $1-h$ it belongs to group L and thus the fact that $x=0$ does not matter. In this state nobody (except bank 1) is aware of the deviation, neither ex ante nor ex post and the payoff must be the same $\mu_L(s)$ as in equilibrium. With probability h it belongs to group H. In that case the planner learns that at least one bank has deviated as the number of high types, $\tilde{N}_H = N_H - 1$, is not N_H as expected. The incentive constraint of bank 1 is therefore

$$(1-p_0)\, h\, (x + \mathbb{E}[\mu_H(s) - \mu_L(s, N_H - 1)]) > p_0 (f(0) - f(x)) \tag{24}$$

Minimizing $\mu_L(s, N_H - 1)$ is good for incentives. If the planner can lower the return μ_L sufficiently, it can implement the first best, i.e., satisfy (24) with $x = x^*$. With limited liability, the first best may not be implementable. The limited liability constraint depends only on the return of bank i, not on N_H. Therefore without loss of generality we can write $\mu_L(s)$ which is either 0 under strict limited liability, or $\mu_L = d/a - s$ under weak limited liability. Once we have minimized μ_L we find the maximum value μ_H^* using the time consistency constraint $h\mu_H^* = \mathcal{M}/A - (1-h)\,\mu_L(s)$ or $h(\mu_H^*(s) - \mu_L(s)) = \mathcal{M}/A - \mu_L(s)$ and the IC constraint (24) becomes

$$(1-p_0)\,(hx + \mathbb{E}[\mathcal{M}(s)/A - \mu_L(s)]) > p_0 (f(0) - f(x)).$$

We know that $\mathcal{M}(s)/A$ is decreasing in γ so it is immediate that the IC improves when γ decreases. We also know that $\mathcal{M}/A$ is decreasing in A. Under strict LL we have $\mu_L = 0$ and know that $\mathcal{M}$ increases with the capital shortfall $K - R = \underline{e}A + \frac{d}{a}A - R$. $\mathcal{M}$ is therefore increasing in d/a and the IC tightens when ex ante leverage is lower. Under weak liability, on the other hand, we have $\mu_L = d/a - s$ and since the slope of $\mathcal{M}$ is less than one we have that $\mathcal{M}(s)/A - \mu_L(s)$ is decreasing in d/a. A lower leverage then loosens the IC constraint.

Proof of Proposition 8. As above the incentive constraint is

$$(1-p_0)\,(hx + \mathbb{E}[\mathcal{M}(s)/A - \mu_L(s)]) > p_0 (f(0) - f(x)).$$

Increasing leverage d/a has an ambiguous effect on incentives because it increases the bailout received by both strong and weak banks.

On the one hand, higher leverage increases the minimal transfer $\mu_L(s) = \max\left\{\frac{d}{a} - r, -\chi r\right\}$:

$$\frac{\partial \mathbb{E}\left[\max\left\{\frac{d}{a} - r, -\chi r\right\}\right]}{\partial\left(d/a\right)} = P\left(\frac{d}{\left(1-\chi\right)a} - hx\right)$$

where $P\left(x\right) = \int_{s\leq x} p_s ds$. Hence the effect of d/a on μ_L is stronger if χ is higher. In the limit $\chi \to 1$, we recover the case of weak LL and $\frac{\partial \mathbb{E}\left[\max\left\{\frac{d}{a} - r, -\chi r\right\}\right]}{\partial\left(d/a\right)} \to 1$.

On the other hand,

$$\frac{\partial\, \mathbb{E}\left[\mathcal{M}(s)/A\right]}{\partial(d/a)} = \mathbb{E}\left[\mathcal{M}'\left(\underline{e}A + \frac{d}{a}A - R\right)\right] = P\left(\frac{\mathcal{K}\left(\gamma\right)}{A} - hx\right).$$

Thus, starting from leverage d/a, locally tightening the leverage constraint (i.e., decreasing d/a) relaxes the incentive constraint if and only if

$$P\left(\frac{d}{\left(1-\chi\right)a} - hx\right) > P\left(\frac{\mathcal{K}\left(\gamma\right)}{A} - hx\right)$$

that is, if the fire sale discount in case of crisis is deep enough:

$$\chi > \hat{\chi} = 1 - \frac{A}{\mathcal{K}\left(\gamma\right)}\frac{d}{a}.$$

Proof of Proposition 9. Let us start from a marginal change. Taking the derivative of (11) we get

$$\frac{dV}{d\alpha_{ij}} = \left(v\left(y_j\right) - v\left(y_i\right) - \left(y_j - y_i\right)v'\left(y_j\right)\right)V'\left(.\right) - \tau$$

This formula contains most of the economics of mergers in our model. If $\tau \to 0$, a marginal asset transfer increases welfare if and only if $y_j > y_i$ and the most attractive first merger is the one between the two furthest banks $i = \arg\min_l y_l$ and $j = \arg\max_l y_l$. This then suggests the following algorithm starting at $k = 0$ with the initial allocation $\{a_i, e_i\}$. Define

$$i^{(k)} \equiv \arg\min_{a_t>0, y_t<0} y_t^{(k)}$$

$$J^{(k)} \equiv \arg\max_t y_t^{(k)}$$

$$j_2^{(k)} \equiv \arg\max_{t\notin J^{(k)}} y_t^{(k)}$$

In words, $i^{(k)}$ is the worst bank among the ones with positive assets and negative surplus (pick any one in case there is a tie), $J^{(k)}$ is the set of best banks, and $j_2^{(k)}$ the next best one. Let $y^c(i,j,\alpha)$ be the capital combination function

$$y^c(i,j,\alpha) = \frac{\alpha_{i,j} y_i + a_j y_j}{\alpha_{i,j} + a_j}$$

The algorithm is then

1. Compute $i^{(k)}, J^{(k)}, j_2^{(k)}$. If $i^{(k)} \in J^{(k)}$, stop. Otherwise proceed.
2. If $y^c\left(i^{(k)}, J^{(k)}, a_{i^{(k)}}^{(k)}\right) > y\left(j_2^{(k)}\right)$ then transfer uniformly all the assets from $i^{(k)}$ to the banks in $J^{(k)}$. Set $a_{i^{(k)}}^{(k+1)} = 0$ and $y\left(J^{(k)}\right) = y^c$. Repeat step 1.
3. Otherwise define α such that $y^c\left(i^{(k)}, J^{(k)}, \alpha\right) = y\left(j_2^{(k)}\right)$, transfer α, set $a_{i^{(k)}}^{(k+1)} = a_{i^{(k)}}^{(k)} - \alpha$ and $y\left(J^{(k)}\right) = y\left(j_2^{(k)}\right)$. Repeat step 1.

It is easy to check that this algorithm provides the welfare $V\left(Av\left(\frac{E-\underline{e}A}{A}\right)\right) - \tau A_-$. The set of failed bank N_- decreases until either $N_- = \emptyset$ or all the banks have the same negative capital surplus. When $E > \underline{e}A$ the algorithm stops when $A_-^{(k)} = 0$, having relocated all failed assets to healthy banks. Capital is not typically equalized across all banks, but since all remaining banks are well capitalized we get $\bar{V}$. When $E < \underline{e}A$ the algorithm does not stop until all the banks have the same capital ratio $\frac{E-\underline{e}A}{A}$. In both cases the algorithm never transfers more than A_-.

Proof of Proposition 10. If $\tau > \tau^*$ and mergers are not used, then the equilibrium payoff of bank 1 under safety x is

$$p_0 f(x) + (1-p_0)\left[h\left(x + s - \frac{d}{a}\right) + (1-h)\,\underline{e}\right]$$

and the only equilibrium features $x_1 = x_2 = \hat{x}$ as in our general Proposition 3. Mergers are too costly and only full bailouts are credible, so we are back into the full moral hazard case with maximal risk-taking.

Contrast this with the case with mergers. If $\tau \leq \tau^*$ then the equilibrium payoff under the first best safety x^* is

$$p_0 f(x^*) + (1-p_0)\left[h^2\left(x^* + s - \frac{d}{a}\right) + h(1-h)\left(x^* + 2\left(s - \frac{d}{a}\right)\right) + (1-h)^2\,\underline{e}\right]$$

The first term in the bracket denotes the expected payoff if both banks succeed, and thus there is no government intervention. The second term denotes the expected payoff if bank 1 succeeds but bank 2 does not, hence bank 1 receives a surplus $s - \frac{d}{a}$ from the merger. The third term captures the case in which both banks get fully bailed out. If the unsuccessful bank 1 is merged to the better bank 2 then the payoff is zero.

If bank 1 instead chooses the minimal safety $x_1 = 0$ then its expected payoff is $p_0 f(0) + (1 - p_0)(1 - h)\underline{e}$. Bank 1 only obtains a positive payoff $\underline{e}$ when bank 2 also fails hence both are bailed out, which happens with probability $1 - h$. Denote Δ_0 the difference between the payoff under minimal safety $x = 0$ and the payoff under first best safety x^*.[11] The incentive compatibility constraint holds if

$$\Delta_0 \leq (1 - p_0)\left[\underline{e}(1 - h(1 - h)) + h(1 - h)\left(s - \frac{d}{a}\right)\right] \tag{25}$$

There exists $\hat{\gamma} > 0$ such that the joint bailout condition (13) and the incentive compatibility constraint (25) both hold for $\gamma < \hat{\gamma}$.

Proof of Lemma 4. Suppose the government has spent $M \geq 0$ in bailout funds with resulting sets of banks N_- and N_+. If $Y_+ = 0$, then no merger takes place and further bailouts of distressed banks happen if $v > \Gamma'(M)$. If $Y_+ > 0$, a merger takes place if there is an i such that $\tau < -y_i \min(v, \Gamma')$.

Consider a bank with $y_i < 0$. If the government does nothing the value is $\mathcal{V}_0 = V_{-i} + vay_i - \Gamma(M)$. If the government bails out the bank by some small amount m the value becomes $\mathcal{V} = V_{-i} + v(ay_i + m) - \Gamma(M) - \Gamma'(M)m = \mathcal{V}_0 + (v - \Gamma'(M))m$ so a (partial) bailout improves welfare if and only if $v > \Gamma'(M)$. Consider a merger instead. Consider next the sale of α of a_i. If $Y_+ = 0$ the acquiring bank has $y_j \leq 0$ so the value becomes $\mathcal{V} = V_{-i} + v(a_i - \alpha)y_i + v\alpha y_i - \tau\alpha - \Gamma(M) = \mathcal{V}_0 - \tau\alpha$. A simple merger reduces welfare. A merger cum recap would lead to $\mathcal{V}_0 + (v - \Gamma'(M))m - \tau\alpha$ which may be positive but is always worse than a simple bailout.
If $Y_+ > 0$ then a merger leads to $\mathcal{V} = V_{-i} + v(a_i - \alpha)y_i - \tau\alpha - \Gamma(M) = \mathcal{V}_0 - v\alpha y_i - \tau\alpha$ which is higher than $\mathcal{V}_0$ if $-vy_i > \tau$. Finally, a full merger of bank i $\alpha = a_i$ is better than a full bailout $m_i = -a_i y_i$ if $-y_i\Gamma'(M) > \tau$.

[11] We have $\Delta_0 = p_0[f(0) - f(x^*)] - (1 - p_0)\left[h\left(x^* + s - \frac{d}{a}\right) - \underline{e}\right] > 0$.

Proof of Proposition 11. Setting a high enough slope δ can achieve the first best: given δ each bank maximizes

$$p_0 f(x_i) + (1-p_0)\,\delta\mathbb{E}\left[r_i\left(1-\frac{1}{N}\right) - \frac{1}{N}\sum_{j\neq i} r_j \mid x_i\right]$$

while the first best safety maximizes

$$p_0 \sum_i f(x_i) + (1-p_0)(1+\gamma)\,\mathbb{E}[R|\mathbf{x}]$$

hence the first best can be implemented using (15) with

$$\delta = \frac{1+\gamma}{1-\frac{1}{N}}. \tag{26}$$

The higher N, the lower is the required δ; when $N=1$, relative performance evaluation cannot help.

To simplify and focus on the core idea we assume that the ex post dispersion in bank returns is small relative to the average return (i.e., idiosyncratic risk is small relative to aggregate risk). To second order in the deviation of returns around the mean we have

$$\sum (e_i + m_i)^{\frac{\eta-1}{\eta}} = N e_*^{\frac{\eta-1}{\eta}} \left(1 - \frac{\eta-1}{\eta} \times \frac{1}{2\eta}\left(\frac{\delta}{e_*}\right)^2 \bar{\sigma}_r^2\right)$$

where $\bar{\sigma}_r = \sqrt{\frac{1}{N}\sum_i (r_i - \bar{r})^2}$ is the standard deviation of returns, equal to the population standard deviation σ_r to first order. Setting a positive slope δ generates a welfare loss relative to the ex post efficient allocation $\Delta V = V\{e_*\} - V\{e_i + m_i\}$ which to second order writes

$$\begin{aligned}\Delta V &= V'\{e_*\}\, N \frac{\eta-1}{2\eta^2} e_*^{1-\frac{1}{\eta}} \left(\frac{\delta}{e_*}\right)^2 \sigma_r^2 \\ &= \frac{N}{2\eta e_*}\delta^2 \sigma_r^2 \gamma\end{aligned}$$

by definition of e_*. Therefore ex post ϵ-efficiency allows to set any slope δ such that $\Delta V \le \epsilon$ or

$$\delta \le \bar{\delta} = \sqrt{\frac{2e_*}{N\gamma\sigma_r^2}\eta\epsilon}. \tag{27}$$

Combining (26) and (27), we find that a sufficient condition to implement the first best is

$$\eta\epsilon \geq \frac{N}{\left(1-\frac{1}{N}\right)^2}\frac{\left(1+\gamma\right)^2\gamma\sigma_r^2}{2e_*}.$$

Note that e_* depends on η, and the dependence can be non-monotone. However, from (14) we have that e_* converges to 0 as $\eta \to 1$ and to some positive constant $k\left(\gamma\right)$ weakly decreasing in γ (solving $V'\left\{k\right\} = \gamma$) as $\eta \to \infty$. Thus for any $\alpha \in (0,1)$ there exists $\eta_\alpha > 1$ such that for $\eta \geq \eta_\alpha$, $e_* > k\left(1-\alpha\right)$ thus we need

$$\eta \geq \max\left\{\eta_\alpha, \frac{1}{\epsilon} \times \frac{N}{\left(1-\frac{1}{N}\right)^2}\frac{\left(1+\gamma\right)^2\gamma\sigma_r^2}{2k\left(\gamma\right)\left(1-\alpha\right)}\right\}.$$

Setting α high enough, the second term dominates, which leads to Proposition 11.

Proof of Proposition 12. We guess and verify that the ex post symmetric allocation $e_1 + m_1 = e_2 + m_2 = e_*$ is not renegotiation-proof, that is $e_* < \bar{e}_1$. Then it must be that the constraint $r_1 - d + m_1 \geq \bar{e}_1$ binds, hence bank 1 gets $\bar{e}_1$ and bank 2 gets $e_2 + m_2$ such that

$$\phi_2\left(\bar{e}_1, e_2 + m_2\right) \times V'\left(\phi\left(\bar{e}_1, e_2 + m_2\right)\right) = \gamma$$

From the renegotiation-proofness principle, we can restrict attention to contracts with $\bar{e}_2 = e_2 + m_2$. Given the return structure, the first best is implementable if $\bar{e}_1, \bar{e}_2$ satisfy:

$$h \cdot \left(\bar{e}_1 - \bar{e}_2\right) = 1 + \gamma$$

where $h = H'(0)$ and H is the c.d.f. of $\epsilon_2 - \epsilon_1$. Therefore $\bar{e}_2 = \bar{e}_1 - \frac{1+\gamma}{h}$. We then look for a solution $\bar{e}_1$ to the equation

$$V'\left(\phi\left(\bar{e}_1, \bar{e}_1 - \frac{1+\gamma}{h}\right)\right) = \frac{\gamma}{\phi_2\left(\bar{e}_1, \bar{e}_1 - \frac{1+\gamma}{h}\right)}.$$

As $\bar{e}_1$ increases from 0 to ∞, the left-hand side decreases from $\lim_{y_2 \to 0} V'\left(\phi\left(\frac{1+\gamma}{h}, y_2\right)\right)$ to 0 and the right-hand side increases from $\lim_{y_2 \to 0} \frac{\gamma}{\phi_2\left(\frac{1+\gamma}{h}, y_2\right)}$ to γ.

Proof of Proposition 13. Suppose that bank i gets $m_i = \frac{\mathcal{K}}{2} + \Delta - r_i$ and bank $j \neq i$ gets $m_j = \frac{\mathcal{K}}{2} - \Delta - r_j$ if and only if $\tilde{\lambda}_i r_i > \tilde{\lambda}_j r_j$ where

$$\tilde{\lambda}_i = \lambda_i + \Lambda_{ji} + \det \Lambda - 1$$

Then $\tilde{\lambda}_1, \tilde{\lambda}_2$ solve the system

$$\begin{aligned} \tilde{\lambda}_1 \Lambda_{11} - \tilde{\lambda}_2 \Lambda_{21} &= \lambda_1 \\ \tilde{\lambda}_2 \Lambda_{22} - \tilde{\lambda}_1 \Lambda_{12} &= \lambda_2 \end{aligned}$$

Therefore

$$\begin{aligned} \mathbb{P}\left[\tilde{\lambda}_1 r_1 > \tilde{\lambda}_2 r_2\right] =& \mathbb{P}\left[\left(\tilde{\lambda}_1 \Lambda_{11} - \tilde{\lambda}_2 \Lambda_{21}\right)(x_1 + s + \epsilon_1) > \left(\tilde{\lambda}_2 \Lambda_{22} - \tilde{\lambda}_1 \Lambda_{12}\right)(x_2 + s + \epsilon_2)\right] \\ =& \mathbb{P}\left[\lambda_1 (x_1 + s + \epsilon_1) > \lambda_2 (x_2 + s + \epsilon_2)\right] \\ =& \mathbb{P}\left[\lambda_1 x_1 - \lambda_2 x_2 > z\right] \end{aligned}$$

where $z = (\lambda_2 - \lambda_1) s + \lambda_2 \epsilon_2 - \lambda_1 \epsilon_1$ has a conditional c.d.f. H. Therefore bank 1's optimal effort x_1 solves

$$\max_{x_1} \quad p_0 f(x_1) + (1 - p_0)\{H(\lambda_1 x_1 - \lambda_2 x_2) 2\Delta\}$$

leading to the first order condition

$$f'(x_1) = \frac{-(1 - p_0)}{p_0} \lambda_1 H'(\lambda_1 x_1 - \lambda_2 x_2) 2\Delta.$$

Similarly, bank 2's optimal effort x_2 solves

$$\max_{x_2} \quad p_0 f(x_2) + (1 - p_0)[1 - H(\lambda_1 x_1 - \lambda_2 x_2)] 2\Delta$$

hence

$$f'(x_2) = \frac{-(1 - p_0)}{p_0} \lambda_2 H'(\lambda_1 x_1 - \lambda_2 x_2) 2\Delta.$$

Therefore, to implement effort levels (x_1^*, x_2^*) that solve $f'(x_i^*) = \frac{-(1-p_0)}{p_0} \lambda_i (1 + \gamma)$ we need

$$\Delta = \frac{1 + \gamma}{2H'(\lambda_1 x_1^* - \lambda_2 x_2^*)}$$

D A Parametric Example

In this section we consider a simple parametric example that can be solved in closed form and illustrates our general results. We also use a case of this model in Figure 4 when studying renegotiation-proof implementation. There are two banks, with sizes $a_1 \geq a_2$. The value function is

$$V(R+M) = \min\left\{0, -\frac{v}{\beta}(K-R-M)^{\beta}\right\}, \quad \beta \geq 1$$

and the cost of funds is linear $\Gamma(M) = \gamma M$. There is only one systemic state, so we omit the s notation. Returns in the systemic state are linear in safety

$$r_i = x_i + \epsilon_i$$

with ϵ_i uniform between 0 and $\bar{\epsilon}$. The normal state return is

$$f(x_i) = -f\frac{x_i^2}{2}$$

Optimal bailout. The optimal bailout in the systemic state is

$$\mathcal{M}(K-R) = \max\left\{0, K-R-\left(\frac{\gamma}{v}\right)^{\beta-1}\right\}$$

Hence the optimized value is

$$\begin{aligned}
\mathcal{V}(R) &= V(R+\mathcal{M}) - \gamma\mathcal{M} \\
&= \begin{cases} -\frac{v}{\beta}\left(\frac{\gamma}{v}\right)^{\beta(\beta-1)} - \gamma\left[K-R-\left(\frac{\gamma}{v}\right)^{\beta-1}\right] & \text{if } R \leq K-\left(\frac{\gamma}{v}\right)^{\beta-1} \\ \min\left\{-\frac{v}{\beta}(K-R)^{\beta}, 0\right\} & \text{otherwise} \end{cases}
\end{aligned}$$

We assume that these returns are low enough that a bailout is always needed in the systemic state: $A(\sup x+1) \leq K - \left(\frac{\gamma}{v}\right)^{\beta-1}$ hence

$$\mathcal{V}(R) = -\frac{v}{\beta}\left(\frac{\gamma}{v}\right)^{\frac{\beta}{\beta-1}} - \gamma\left[K-R-\left(\frac{\gamma}{v}\right)^{\beta-1}\right].$$

First Best. The first best safety x^* is the same for both banks and solves

$$\begin{aligned} x^* &= \arg\max_x p_0 A f(x) + (1-p_0)\left(Ax + \mathbb{E}\left[\mathcal{V}(R)\,|x\right]\right) \\ &= \arg\max_x p_0 A f(x) + A(1-p_0)(1+\gamma)\,x \end{aligned}$$

hence

$$x^* = \frac{q(1+\gamma)}{f}$$

where $q = \frac{1-p_0}{p_0}$ is the odds ratio of a crisis. x^* is increasing in q and increasing in γ.

Moral hazard with symmetric bailouts. Suppose bailouts are proportional to bank size:

$$m_i = \frac{a_i}{A}\mathcal{M}(K-R).$$

Then bank i solves

$$\hat{x}_i = \arg\max_x p_0 a_i f(x_i) + (1-p_0)\left(a_i x_i + \frac{a_i}{A}\underbrace{\left[K - a_i x_i - a_j x_j - \left(\frac{\gamma}{v}\right)^{\beta-1}\right]}_{=\mathbb{E}[\mathcal{M}(K-R)|\mathbf{x}]}\right)$$

Thus

$$\hat{x}_i = \frac{q}{f}\left(1-\frac{a_i}{A}\right) < x_i^*.$$

With symmetric bailouts, both banks take excessive risk, and the moral hazard problem is worse for the larger bank (high a_i/A). This is consistent with Dávila and Walther (2020)'s results on symmetric bailouts with small and large banks.

Tournament with bonus-malus. With symmetric banks $a_i = a$, the credible tournament described in Section 3 with

$$\Delta = \frac{1}{2}a\bar{\epsilon}\left(\gamma + \frac{1}{2}\right)$$

implements the first best safety. Δ is increasing in $\bar{\epsilon}$: noisier returns require larger rewards.

With asymmetric banks, under the condition

$$\frac{a_1}{a_2}\left(\frac{\frac{a_1}{A}+\gamma}{\frac{a_2}{A}+\gamma}\right)\leq 1+\frac{\bar{\epsilon}}{x^*}=1+\frac{\bar{\epsilon}f}{q\left(1+\gamma\right)}$$

the handicapped tournament (22) with

$$\lambda=\frac{a_1}{a_2}\left(\frac{\frac{a_1}{A}+\gamma}{\frac{a_2}{A}+\gamma}\right)$$
$$\Delta=\frac{\frac{1}{2}a_1\bar{\epsilon}\left(\gamma+\frac{a_1}{A}\right)}{1-\left(\lambda-1\right)x^*/\bar{\epsilon}}$$

implements the first best safety.

Limited liability. As in the main text we consider a tournament rule that satisfies strong limited liability by transferring the total bailout $\mathcal{M}$ to bank 1 if $\lambda r_1\geq r_2$ and to bank 2 otherwise. Bank 1 solves

$$\max_{x_1} p_0a_1f\left(x_1\right)+\left(1-p_0\right)\left(a_1x_1+\left[K-a_1x_1-a_2x_2-\left(\frac{\gamma}{v}\right)^{\beta-1}\right]\int_0^{\bar{\epsilon}}G_\epsilon\left(\lambda\epsilon_1+\lambda x_1-x_2\right)d\epsilon_1\right)$$

where G_ϵ is the c.d.f. of ϵ_1. With $a_1=a_2=a$ and $\lambda=1$ the maximal implementable safety $x^{\max}$ satisfies

$$p_0f'\left(x^{\max}\right)+\left(1-p_0\right)\left[\frac{1}{2}+\frac{K-Ax^{\max}-\left(\frac{\gamma}{v}\right)^{\beta-1}}{a}\min\left\{1,\frac{1}{\bar{\epsilon}}\right\}\right]=0$$

or

$$x^{\max}=\frac{q}{f+2q\min\left\{1,\frac{1}{\bar{\epsilon}}\right\}}\left[\frac{1}{2}+\frac{K-\left(\frac{\gamma}{v}\right)^{\beta-1}}{a}\min\left\{1,\frac{1}{\bar{\epsilon}}\right\}\right]$$

which is indeed decreasing in γ (and above x^* for γ low enough) and in bank size a.