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

Thomas Philippon[†] Olivier Wang[‡]

September 19, 2022

Abstract

We study time-consistent bank resolution mechanisms. The key constraint is that governments cannot avoid bailouts that are ex post efficient. 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 incentives. We analyze properties of credible tournament mechanisms that provide support to the best performing banks and resolve the worst performing ones. We extend our mechanism and show that it continues to perform well when banks are heterogeneous in size, when they are imperfect substitutes, and when they are differentially interconnected as long as bailout funds can be earmarked.

*We are grateful to Andrei Shleifer and three anonymous referees, 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 on 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. When the limited liability constraint binds, our model offers a macro-prudential justification 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.

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. 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 ex post welfare deviates by less than ϵ from its 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.

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 now present our baseline environment before defining the first best allocation. Moral hazard appears because banks anticipate government support policies when deciding how much risk to take.

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 that maps realized returns on banks' assets to government transfers. Each bank then chooses a safety investment $x_i \in [0, \bar{x}]$. Uncertainty, resolved at time $t = 1$, 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_i and deposits with face value d_i due at time 1. We denote by r_i^s the gross asset return of bank i in state s at time 1. The equity of the bank is $e_{i,s} = a_i r_i^s - d_i$ before any intervention. We say that a bank is well capitalized ex post when $e_{i,s} \geq \kappa a_i$ or equivalently $r_i^s \geq \underline{r}_i = d_i/a_i + \kappa$, and its capital surplus is then $e_{i,s} - \kappa a_i$. The equity of the bank is $e_{i,s} + m_{i,s}$ after intervention where $m_{i,s}$ is the cash injection from the government. The variable $m_{i,s}$ 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.[1]

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 $[0, \bar{x}]$ 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

[1] 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.

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

times. The maximal risk banks can take, $x = 0$, leads to the 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., featuring no crisis and no bailout. The government's value function in state s

$$V\left(\{e_{i,s} + m_{i,s}\}_{i=1..N}\right)$$

is concave and weakly increasing in each argument $e_{i,s} + m_{i,s}$. To simplify the notation we often write $V\{e_i + m_i\}$.

V is flat at its maximum when all banks are well capitalized: $V = \bar{V}$ when $e_i \geq \kappa a_i$ 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 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 $M_s = \sum_i m_{i,s}$ 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 $R = \sum_i r_{i,s}$ is the random aggregate asset return.[3]

For simplicity we will first consider the case where all the banks are identical ex ante: $a_i = 1$ and $d_i = d$ for all i. We wish to focus our analysis on the issue of undercapitalization during systemic crises, not on the pricing of deposit insurance. We

[3] 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).

therefore assume there is no default on debt d:

Assumption 2. $d \leq \min\{r_{i,s}\} < \underline{r} = d + \kappa$.

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 micro-foundations in Appendix B. Broadly speaking, two classes of models deliver the welfare function specified above. The first class includes models of runs such as Diamond and Rajan (2012). A bank with low equity (but still potentially solvent) faces the risk of a run unless it restructures part of its debt; restructuring, however, can trigger money market disturbances, including further runs as happened after the collapse of Lehman Brothers in 2008. The second class includes models of credit crunch (Myers, 1977; Holmström and Tirole, 1997; Philippon and Schnabl, 2013). In these models, new investment opportunities arise at date-1, but limited pledgeability or debt overhang prevents solvent banks from investing efficiently 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.[4]

Assumption A2 means that capital requirements or more generally TLAC requirements are calibrated so as to protect small depositors, where 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 equity, and we lump all assets returns into one category that we call gross value, or output.[5]

Adding defaultable debt (i.e., removing Assumption A2) would make the model more complex but yield very similar results. With default, bailouts $\{m_{i,s}\}$ reduce banks' ex ante funding cost on risky debt by reducing the probability of default and raising the

[4]One advantage of using a welfare function V is to highlight a key feature that is not typically discussed in micro-founded models. As our analysis makes clear, the critical feature determining the performance of our mechanism 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.

[5]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.

recovery value. The lower funding cost allows banks to reduce safety x_i and scale up risky investments paying off in the normal state, as captured by the decreasing function $f(x_i)$ in our setup. This would be true whether x_i is set before or after debt is priced: in the first case x_i and the resulting credit risk are specified in the debt contract, and in the second case bank creditors price the debt based on rational expectations about the bank's optimal choice of x_i, which in turn depends on the bailout policy. Moreover, in a standard setting featuring creditors subject to a participation constraint, the ultimate benefit from government guarantees would still accrue to equityholders (who set x_i) just like in our model.

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.

1.2 No Bailouts

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 $r_{i,s}$. Let $\tilde{x}$ be the autarky safety, that is the privately optimal safe return of a bank anticipating $m = 0$ in all states:

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

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.

Consider next the socially optimal allocation when there are no bailouts. 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

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

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 A3 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 $f\left(x^*\right) > \underline{\mathrm{r}}$ since f is decreasing).

The program of the planner is therefore

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

We define the ex post optimal vector of bailouts as

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

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 bailouts with a strictly positive probability, but rather what the anticipation of a bailout does to private incentives for safety.

2 Credible Tournaments

In the main text we focus on the following special case of the model that illustrates our results in the simplest possible form. Appendix C studies the general case.

Setup. There are two banks $N = 2$ with identical sizes $a_i = 1$ and two aggregate states: a normal state with probability p_0 and a crisis state with probability $1 - p_0$. Bank returns depend on an ex ante safety investment $x_i \in [0, \bar{x}]$. The normal state return is decreasing in safety: $f(x_i) = \bar{r} - f\frac{x_i^2}{2}$. The return in the crisis state is increasing in safety: $r_i = x_i + \epsilon_i$, where the idiosyncratic risk ϵ_i is distributed uniformly between 0 and $\bar{\epsilon}$ and independent across banks.

Before any government intervention bank i has equity $e_i = r_i - d$, where d is debt. The government can intervene in the crisis state by injecting a net transfer or "bailout" m_i so that equity becomes $e_i + m_i$. The shadow cost of public transfers is linear: $\Gamma(M; \gamma) = \gamma M$ hence a country with lower γ has more fiscal space.

Social Welfare and First Best Allocation. The social planner chooses x_i and m_i to maximize

$$\mathbb{E}\left[\sum_i e_i + V\Big(\sum_i (e_i + m_i)\Big) - \gamma \sum_i m_i | \boldsymbol{x}\right].$$

We define the aggregate capital requirement as $\kappa = \sum_i a_i \kappa$ which is simply $\kappa = 2\,\kappa$ in the model with two identical banks. The function

$$V(E) = \min\left\{0, -\frac{v}{2}(\kappa - E)^2\right\} \tag{5}$$

captures the externalities imposed by distressed banks. We assume that the severity of the crisis is such that a bailout is always needed in the systemic state but never in the normal state.[6]

Importantly, we assume in this benchmark that the externality V only depends on the aggregate health of the banking sector $E = \sum_i e_i$. We relax this "pure systemic

[6]The condition is $\bar{x} + \bar{\epsilon} + \frac{\gamma}{2v} \leq \kappa + d \leq \bar{r} - f\bar{x}^2/2$.

risk" assumption in later sections, but view it as a good starting point to capture the deadweight loss from an undercapitalized banking system.

The first best allocation can be solved in two steps. Ex post, the optimal aggregate bailout in the systemic crisis state is

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

where $K=2\left(d+\kappa\right)$. $\mathcal{M}$ decreases with γ: fiscal slack allows for a larger bailout. In the limit of costless bailouts $\gamma\rightarrow 0$, the government never lets the aggregate capitalization E fall below κ. With a positive cost γ, the government allows some aggregate undercapitalization, up to a threshold γ/v.

Ex ante, the first best safety x^* is

$$x^*=\frac{q}{f}\left(1+\gamma\right) \tag{6}$$

where $q=\frac{1-p_0}{p_0}$ is the odds ratio of a crisis. x^* is increasing in q and γ: efficiency requires more self-insurance by banks if a crisis is more likely and when government insurance is more expensive. By contrast, the no-bailout privately optimal safety defined in (3) is

$$\tilde{x}=\frac{q}{f}$$

and ignores the externality from V or equivalently the fiscal externality captured by γ.

Equilibria under Limited Commitment. We next consider equilibria under different policy regimes. Expectations about the policy rule for m_i affect the private ex ante choices of safety x_i. The key friction is that the government lacks commitment and cannot commit not to intervene. As a result, banks always correctly expect transfers m_i to satisfy the *credibility constraint*

$$\sum_i m_i=\mathcal{M}\left(K-R\right) \tag{7}$$

for any realization of returns.

2.1 Moral Hazard under Symmetric Bailouts

We start by showing the moral hazard problem that arises when the government lacks commitment *and* bailouts are symmetric across banks, as discussed by the existing literature. Suppose that each bank gets half of the aggregate bailout:

$$m_i = \frac{\mathcal{M}\left(K - R\right)}{2}.$$

Then bank i sets x_i to maximize

$$p_0 f\left(x_i\right) + \left(1 - p_0\right)\left(x_i + \frac{1}{2}\left[K - x_i - x_j - \frac{\gamma}{v}\right]\right). \tag{8}$$

The equilibrium safety is

$$\hat{x} = \frac{q}{2f}.$$

With symmetric bailouts, both banks take excessive risk. More precisely, we have

$$\hat{x} < \tilde{x} \le x^*,$$

that is, $\hat{x}$ departs from the first-best safety x^* in two ways. First, there is "collective moral hazard" (Farhi and Tirole, 2012): each of the two banks realizes that it is insured against half of its risk-taking by the government and thus chooses a safety that is only half of the no-bailout choice $\tilde{x} = q/f$. Second, the no-bailout safety $\tilde{x}$ is itself lower than the first best safety x^* that takes into account the fiscal externality and thus increases with γ. Next, we propose a mechanism that solves both problems.

Heterogeneous banks. More generally, with N banks of size a_i such that $A = \sum_i a_i$, and symmetric bailouts $m_i = \frac{a_i}{A}\mathcal{M}\left(K - R\right)$, the first best safety would still be $x^* = \frac{q}{f}(1+\gamma)$ for all banks. However the equilibrium safety of bank i would be $\hat{x}_i = \frac{q}{f}\left(1 - \frac{a_i}{A}\right)$, making the moral hazard problem worse for larger banks, consistent with Dávila and Walther (2020)'s results on symmetric bailouts with small and large banks.

More crisis states. In this simple setup the equilibrium safety $\hat{x}$ does not depend on γ because there is only one crisis state, which is severe enough that the probability of bailout is always 1 in that state. With more crisis states s as in our general setup in Appendix C, fiscal space also affects the probability of bailout and thus $\hat{x}$ is increasing in γ: banks invest less in safety if fiscal capacity is high (low γ) because the government provides insurance against more realizations of the aggregate shock.

2.2 First Best under Tournaments

We now analyze a mechanism relying on *asymmetric* bailouts. Consider the following tournament rule $\mathscr{T}$ that sets:[7]

$$m_i = \begin{cases} \frac{\mathcal{M}(K-R)}{2} + \Delta & r_i > r_j \\ \frac{\mathcal{M}(K-R)}{2} - \Delta & r_i < r_j \end{cases}$$

Instead of injecting the same amount of equity in both banks, this rule introduces a wedge $\Delta \geq 0$. The bank with the higher realized return obtains a higher bailout. By construction this rule is credible, that is, the aggregate bailout satisfies the time-consistency constraint (7) in all states of the world. Bank i sets x_i to maximize

$$p_0 f(x_i) + (1-p_0)\left(x_i + \frac{1}{2}\left[K - x_i - x_j - \frac{\gamma}{v}\right]\right) + 2\Delta\mathbf{P}\left[r_i > r_j | x_i, x_j\right]. \tag{9}$$

Our main result is that in stark contrast the case of symmetric bailouts, this tournament mechanism can substantially mitigate moral hazard, and even implement the first best with the appropriate Δ:

Proposition 1. *The tournament mechanism $\mathscr{T}$ with*

$$\Delta^* = \frac{1}{2}\bar{\epsilon}\left(\gamma + \frac{1}{2}\right) \tag{10}$$

implements the first best safety $x_1 = x_2 = x^*$.

The objective function (8) under symmetric bailouts corresponds to a wedge $\Delta = 0$. Moral hazard arises because the term $\frac{1}{2}\left[K - x_i - x_j - \frac{\gamma}{v}\right]$ is decreasing in x_i: from each bank's perspective, investing in safety has the downside of decreasing the aggregate bailout $\mathcal{M}$. Under the tournament mechanism, banks' objective function (9) contains an additional term $2\Delta\mathbf{P}\left[r_i > r_j | x_i, x_j\right]$. The crucial intuition is that this term is increasing in x_i and it makes bank i's objective function *supermodular* in (x_i, Δ). Therefore a higher wedge Δ leads banks to choose higher safety, and this force can be strong enough to counteract the moral hazard term completely by setting the right Δ. In Appendix C we use a much more general framework that highlights the role of supermodularity.

[7] Each bank gets $\frac{\mathcal{M}(K-R)}{2}$ in case of tie $r_i = r_j$, which is a zero probability event.

Comparative Statics. Equation (10) provides a transparent closed-form for the optimal wedge Δ^* and reveals its main determinants. Δ^* is increasing in $\bar{\epsilon}$, which captures the magnitude of idiosyncratic risk: noisier bank-specific returns require larger rewards. This is a standard result of incentive models (Holmström, 1979; Lazear and Rosen, 1981). On the other hand, and perhaps surprisingly, aggregate risk is irrelevant: the odds ratio of a crisis $q = \frac{1-p_0}{p_0}$ does *not* appear in Δ^*. This is important because it implies that even when crises are unlikely the required Δ^* may not be large. While it is true that a low crisis probability weakens the incentive effect of a given wedge Δ, a lower q also weakens the severity of moral hazard in the first place. These two forces cancel out exactly, making the optimal wedge Δ^* independent of q. This also highlights the key point that the benefit from the tournament mechanism is not to prevent the crisis altogether but to neutralize the moral hazard component. Section 4 extends the framework to allow for neglected risk.[8]

Finally, the optimal wedge Δ^* also increases with γ: countries with less fiscal space (higher γ) need to discriminate more between good and bad performers. This is because the required investment in safety is higher when there is less flexibility to intervene ex post. In this case solving the moral hazard becomes even more crucial, which justifies setting a higher wedge.

Interpretation of the wedge Δ. 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 welcoming 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.

While we agree that convincing healthy banks to participate in programs of equity injection can be difficult, we want to emphasize some crucial features of our model that differ from existing work. First, accepting public support is a sign of weakness in standard mechanisms *because* they provide more support to weaker institutions. In tournament mechanisms, however, public support is a signal of strength. In fact, with our mechanism, when one bank is allowed to fail, the market value of the other ones should *increase* because they are now more likely to benefit from government support. Second, we note that all banks, including the best capitalized ones, seem to welcome

[8] In the current setup there is only one crisis state, so $1 - p_0$ has a straightforward interpretation as the probability of a crisis. More generally, in Appendix C we show that the optimal wedge Δ^* is still independent of p_0, but it can depend on the relative likelihood of the different crisis states $s \neq 0$.

subsidized mergers with asset guarantees. This is consistent with a reverse stigma where market participants know that the government would select strong banks to take over weaker ones. An important result of our paper (see Section 3.2.2) is that mergers indeed provide tournament-like incentives.

Beyond stigma, another reason preventing take-up by some banks was that government bailouts can feature rather punitive terms. In the language of our model, these conditions are designed to minimize the ex post fiscal cost $\Gamma(M;\gamma)$ as in Philippon and Schnabl (2013). For instance, during the 2007-2009 crisis, the Capital Purchase Program included restrictions on common stock dividends and executive compensation. Government equity took the form of preferred stock and missed dividend payments led to appointment of board directors by the Treasury, which banks actively tried to avoid (e.g., Mücke et al. 2022). We come back to this point in Section 5.

3 Limits of Tournaments

In this section we extend our framework along several dimensions. We show how to adapt our tournament mechanism but also highlight the limits of the mechanism. In particular, 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 Sections 3.2 and G, is that the planner might not be *willing* to punish because of imperfect substitutability between banks or financial contagion.

3.1 Limited Punishments

The previous scheme is attractive in its simplicity, but may run against a limited liability constraint if the required wedge Δ is high. Let us now consider the case where government transfers and taxes are constrained by a limited liability (LL) constraint:

$$m_{i,s} \geq 0.$$

Consider the following alternative tournament rule $\mathscr{T}_{LL}$ that transfers the total bailout $\mathcal{M}$ to bank 1 if $r_1 > r_2$ and to bank 2 otherwise:[9]

$$m_i = \begin{cases} \mathcal{M}\,(K-R) & r_i > r_j \\ 0 & r_i < r_j \end{cases}$$

The rule $\mathscr{T}_{LL}$ satisfies strong limited liability by construction.

Proposition 2. *The highest safety implementable under limited liability is given by*

$$x^{\max} = \frac{q}{f+2q}\left[\frac{1}{2} + K - \frac{\gamma}{v}\right]$$

and is decreasing in the shadow cost of public funds γ.

Proposition 2 gives a striking result with respect to fiscal slack: a lower cost γ *increases* safety. This is exactly the opposite of the conventional wisdom based on symmetric mechanisms. With symmetric bailouts, fiscal slack implies insurance against more systemic states and thus a more acute moral hazard problem. With asymmetric bailouts, fiscal slack gives the government more flexibility to reward the winners of the tournament. This reward improves incentives when harsh punishments for the tournament's losers are not feasible.

Our framework gives a macro-prudential reason for clawback provisions on executive compensation as they help relax the binding limited liability constraint. One should also 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.

Remark 1. There are two ways to write limited liability. We studied the strict form ("strict LL") that imposes non-negative net transfers $m_i \geq 0$. This constraint typically leaves equity holders with a surplus. A weaker form of limited liability ("weak LL") is $e_i + m_i \geq 0$, which allows negative transfers of residual equity value, but not more. Since punishments can be higher under weak LL, incentives are naturally stronger. In Appendix E 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.

[9] As before $\mathcal{M}$ is split equally in case of tie $r_i = r_j$.

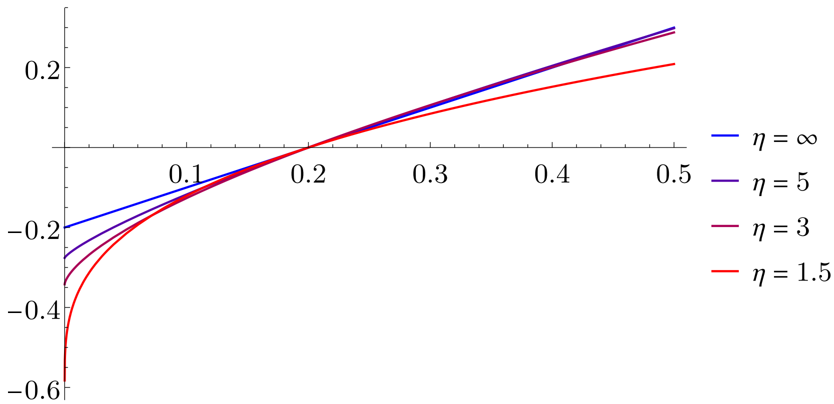

Figure 1: Function $V(e) = e^{\frac{\eta-1}{\eta}} - \kappa^{\frac{\eta-1}{\eta}}$ for different values of η. Lower η makes the function more concave, which increases incentives to offset individual capital shortfalls.

3.2 Differentiated Banks

The "pure systemic risk" model considered thus far supposes a value function V that only depends on the aggregate capital of the banking sector. This fungibility 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. (2017). Suppose then that there are N imperfectly substitutable banks and the value function is

$$V\{e_i + m_i\} = V(\phi\{e_i + m_i\} - \phi\{\kappa\}), \quad \text{where } \phi\{e_i\} = \sum_{i=1}^{N} e_i^{\frac{\eta-1}{\eta}}. \tag{11}$$

ϕ is a constant elasticity of substitution (CES) aggregator and $\eta > 1$ is the elasticity of substitution between banks. This value function converges to the one in the pure systemic model (18) as $\eta \to \infty$. It also captures the fact that it becomes more costly to take away the positive equity e_i from bank i as it gets smaller.

Without commitment, perfect ex post efficiency requires equalizing the marginal return of transfers m_i across banks i, that is for each i

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

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 $e_* > \kappa$

solves

$$\frac{\eta - 1}{\eta} e_*^{\frac{-1}{\eta}} V' \{e_*\} = \gamma \tag{12}$$

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

At first glance, it seems that imperfect substitutability brings back an extreme form of moral hazard. 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. But 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.

3.2.1 Partial Commitment

We now relax the assumption of complete lack of commitment and re-establish our main result by introducing a small amount of commitment. We give 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. Formally, for any realization $\{e_i\}$ the government can choose transfers $\{m_i\}$ such that

$$\left| V\{e_i + m_i\} - \gamma \sum_i m_i - \max_{\{m_i\}} \left[V\{e_i + m_i\} - \gamma \sum_i m_i \right] \right| \leq \epsilon.$$

The case $\epsilon = 0$ is the standard time-consistent implementation. The next proposition shows that the complete lack of commitment is a knife-edge outcome. 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.[10]

Consider a mechanism that transfers

$$m_i = e_* + d - r_i + \delta \left(r_i - \bar{r} \right) \tag{13}$$

[10] In Appendix F we consider a second, and independent, relaxation of the notion of time-consistency. We analyze renegotiation-proof mechanisms (Fudenberg and Tirole, 1990): the government can only deviate from promises if this generates a Pareto-improvement. This solution concept provides a weak form of commitment consistent with the political economy of bailouts. The idea is that ex post discretionary policies are less likely to generate backlash and intense lobbying if all the involved parties (the government and the different banks) benefit. We show that under this relaxation tournaments can again implement the first best with sufficient fiscal capacity.

to each bank so that the capital after bailout is $r_i - d + m_i = e_* + \delta(r_i - \bar{r})$ where e_* is the ex post efficient (symmetric) capital that solves (12) and $\bar{r} = \frac{1}{N}\sum_i r_i$ is the average return. This relative performance evaluation mechanism is in the spirit of tournaments, but slightly simpler to use here. 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.

Proposition 3. *There exists $\alpha \in (0,1)$ increasing in ϵ such that the first best is implementable under ϵ-commitment using transfers*

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

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

$$\eta\epsilon \geq \frac{N}{\left(1-\frac{1}{N}\right)^2} \frac{(1+\gamma)^2 \gamma \sigma_r^2}{2k(\gamma)(1-\alpha)}. \tag{14}$$

The right-hand side of (14) is increasing in γ, in the variance of returns σ_r^2, and in the number of banks N.

Equation (14) 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 number 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 3 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.

3.2.2 Mergers

For simplicity we formalized the implementation of the tournament policies using only 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. We now extend the model by adding a *resolution authority*, that is a technology that allows the government to write for any undercapitalized bank's $e_i < \kappa$ capital claims to 0 and transfer the assets and deposits to another bank or another set of banks, at some cost τ per unit of assets.

As before we adopt the value function (11), focusing on $N = 2$ banks with equal size $a = 1$ and a linear value $V(x) = x$ (conditional on banks being undercapitalized). Thus

$$V\{e_i\} = \sum_i v(e_i)$$

where $v(e_i) = e_i^{\frac{\eta-1}{\eta}}$. As shown by (12), simple bailouts without mergers lead to full moral hazard: the government bails out both banks in case of crisis to replenish their post-injection equity $e_i + m_i$ to $e_* = v'^{-1}(\gamma)$, and as a result both banks are fully insured and choose the minimal safety $x_i = 0$.

Suppose now that ex post the government can decide to merge the bank with the lower realized return, say bank 2, with bank 1, and then recapitalize the merged entity. The optimal post-merger bailout is $M = E_* - e_1 - e_2$ where $E_* = v'^{-1}(\gamma/2)$. This sequence of interventions yields a final value

$$V^{\text{post}}\{e_i\} = 2\left(v\left(E_*\right) - v(\kappa)\right) - \gamma\left(E_* - e_1 - e_2\right) - \tau$$

As a result the merger followed by a bailout to the merged entity dominates the simple bailouts without mergers if the merger cost τ is low enough:

$$\tau \leq \tau^* (\gamma) \tag{15}$$

where τ^* is a decreasing function.[11] Therefore under condition (15) bank i's shareholders anticipate ending up with equity E_* if $r_i > r_j$, and 0 otherwise, which gives powerful incentives to invest in safety, just like in a tournament that uses asymmetric transfers.

Proposition 4. *If $\tau \leq \tau^*(\gamma)$ then mergers are optimal and the equilibrium safety $\hat{x}$ is $\hat{x} = \frac{q}{f} v'^{-1} (\gamma/2)$.*

This model extension makes a simple point: mergers work in the right direction for incentives and they are especially useful with imperfect substitutable banks. Our current analysis is just a first pass and abstracts from two key concerns. First, the U.S. banking sector is already quite consolidated at this point. Our framework suggests that in addition to the usual harmful effects on competition in normal times, this concentration at the top also undermines the virtuous incentive effects created by the threat of mergers in bad times. One solution would be to refine the merger process by breaking up the target bank and sell its divisions to several other banks. Second, the incentive effects of mergers come from wiping out the shareholders of the acquired bank. In the paper we treat shareholders and managers as a single entity, but if operational efficiency requires maintaining the management of the target, the ex ante incentives of managers may be weakened. Our implicit assumption is that other contracts can be written within each bank to align the incentives of managers with those of their shareholders. Part of the merger cost τ can be viewed as capturing the associated costs, for instance if the merger involves firing a fraction of managers even though this entails an efficiency loss. In Philippon and Wang (2022) we study in depth the interactions between ex ante incentives and ex post merger and bailout policies in a more general environment.

3.3 Financial Contagion

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

[11] The full expression is $\tau^* (\gamma) = 2 \left[v \left(v'^{-1} (\gamma) \right) - v \left(v'^{-1} (\gamma) \right) \right] - \gamma \left(v'^{-1} (\gamma/2) - 2 v'^{-1} (\gamma) \right)$. For instance, for $\eta = 2$ we have $v(x) = \sqrt{x}$ and $\tau^* = \frac{3}{4\gamma}$.

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.[12]

[XXX General framework in Appendix]

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

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

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.

To illustrate the result, suppose $N = 2$ and 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 (28) 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.

The takeaway from this section is that financial contagion undermines credibility if and only if bailout funds can flow freely 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).

However, if given a choice ex post the government would prefer contagious bailouts (29). This motivates ex ante measures that limit the options ex post and put the government in a position to credibly earmark the bailout funds. For instance, our model sheds new light on the "safe harbor" versus "automatic stay" debate.[13] 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 safe harbor provisions) and macro regulation (systemic risk management under limited commitment).

4 Moral Hazard, Neglected Risk, and Macroprudential Policy

[XXX ABOUT MORAL HAZARD]

We do not claim that moral hazard was the leading factor in the 2008 crisis. Multiple factors. Mounting evidence for overoptimism, to which we come back below. But moral hazard has been at the heart of the policy debate (SOURCES) and we argue that even under the hypotheses that have provided the best case for moral hazard, a resolution regime based on tournaments can go a long way in mitigating the risk-taking.

Second, our approach can be extended to discuss the alternative main driver overoptimism.

4.1 Neglected Risk

Our model assumes that banks and regulators hold correct and thus identical beliefs about the risk of a systemic crisis. How does our tournament mechanism interact with overoptimism, a major driver of bank risk-taking (Gennaioli et al. 2015, Baron and Xiong 2017)?

[13] 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.

Realist Government Suppose that banks hold potentially wrong beliefs, captured by the likelihood ratio $q_i = \frac{1-p_0^i}{p_0^i} > 0$ of the aggregate crisis state, while the government/social planner has a correct belief q. To focus on the role of beliefs we assume costless bailouts, $\gamma = 0$. From the perspective of the planner, the first best allocation still features the same safety as before $x_i = x^* = q/f$ by all banks.

[XXX INFORMATION STRUCTURE: Agree to disagree]

As before, we compare equilibria under limited commitment for different resolution regimes. With symmetric bailouts, the equilibrium safety is

$$\hat{x}_i = \frac{1}{2}\frac{q_i}{f}.$$

Relative to the rational expectation case, banks can now take excessive risk for two reasons: moral hazard for given beliefs (captured by the factor 1/2 in the case of two banks), but also overoptimism ($q_i < q$). Moreover, if banks hold heterogeneous beliefs, the more optimistic ones will take more risk.

Consider instead the tournament rule $\mathscr{T}$ with some wedge $\Delta > 0$. The credibility constraint (7) implies that the wedge Δ must be the same for all banks, and thus the equilibrium safety is

$$\hat{x}_i = \frac{q_i}{f}\left(\frac{1}{2} + \frac{2\Delta}{\bar{\epsilon}}\right).$$

Therefore the tournament rule can implement an average safety $\frac{1}{2}(x_1 + x_2) = x^*$ (and the first best allocation $x_1 = x_2 = x^*$ if both banks are equally optimistic) by setting a wedge

$$\Delta^* = \frac{\bar{\epsilon}}{2}\left(\frac{1}{2} + \frac{q - q^a}{q^a}\right)$$

where $q^a = \frac{1}{2}(q_1 + q_2)$ is the average private belief. Relative to expression (10) under rational expectations, there is an additional correction from the term $\frac{q-q^a}{q^a}$. The more optimistic banks are relative to the planner, the higher wedge Δ^* is needed to induce investments in safety.

Optimist Government

4.2 Macroprudential Policy and Tournaments

Heterogeneous banks: optimistic and rational

Need a macropru leakage Hanson et al. (2011), Plantin (2014), Farhi and Tirole (2020), Bengui and Bianchi (2022)

Suppose some subset of banks can evade $x \geq \underline{x}$ constraint?

5 Discussion and 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.

We conclude with a discussion of some practical issues in the implementation of tournament-like incentives. For theoretical clarity we have analyzed rewards and punishments as taxes and transfers but it is useful to understand the political and economic forces that may lead to outcomes similar to those described above. Policymakers may want to lean into, rather than resist such forces.

The Bear Stearns - Lehman Brothers - AIG Sequence A useful way to discuss implementation is to ask how our mechanism would have played out in September 2008. We study two dimensions – decisions by government officials, and reactions by market participants – and we ask if they can be rationalized within the model.

The most dramatic sequence of the Great Financial Crisis is the failure of Lehman Brothers followed by the bailout of AIG and Money Market Mutual Funds. This sequence is consistent with our model if we interpret failure as a punishment and the bailout as a way to stabilize the financial system. The equilibrium revealed by the reaction of market participants, however, is not consistent with the prescription of our model. In our model, when one bank is allowed to fail, the market values of the other ones *increase* because they are now more likely to benefit from government support. In our framework participants would interpret the failure of Lehman as a necessary step to avoid moral hazard, but once this was done, they would price in *more* support for the rest of the system. In reality market participants interpreted the bankruptcy as a signal of *less* support in the future, or at least more uncertainty as to whether support would be forthcoming.

[XXX MORE DETAILS HERE? ADD EXTENSION IN APPENDIX?]

The Bear Stearns-Lehman sequence is also partly consistent and partly inconsistent with our model. The acquisition of Bear Stearns by JP Morgan – including the use of asset guarantees – is clearly consistent with our model. On the other hand the sequence between the successful sale of Bear Stearns and the failed attempts to sell Lehman Brothers (to Bank of America, the Korean Development Bank, and Barclays) is inconsistent with our model. Former Treasury Secretary Paulson describes how Lehman Brothers's CEO Richard Fuld interpreted the terms of the previous sale: "*Dick* [Fuld] *did not want to consider any offer below $10 per share. Bear Stearns had gotten that, and he would accept nothing less for Lehman.*" (Paulson, 2010, p. 173)

These two examples highlight the gap between the expectations of agents in our model and those of investors and participants during the 2008 crisis. A plausible explanation is that agents in our model know how to interpret the actions of the government. They understand that the government lets some banks fail to provide incentives but maintains its commitment to stabilize the system. This policy, however, was not spelled out explicitly and was not understood by market participants.

Runs, Arbitrary Decisions, and Incentives Policy making during financial crises requires real-time decisions with limited information and under political constraints. It is no surprise, then, that some decisions appear poorly motivated. This seeming arbitrariness, however, can be consistent with our model. With optimal incentives under moral hazard (Holmström, 1979) all agents take the same incentive-compatible action: any ex post difference in outcomes is purely random. Rewards and punishments are thus literally arbitrary *in equilibrium.* In our model, when banks are symmetric ex ante they all make the same investment in safety. Good performance *in equilibrium* reflects good luck, and bad performance bad luck. The fact that banks are rewarded for luck is a feature of the equilibrium.

Incentives arise from the increasing the likelihood of punishment if a bank deviates from the prescribed level of safety. Arbitrariness is detrimental because it decreases the sensitivity of performance to action. The noise component ϵ captures this effect in our model. An increase in the variance of ϵ requires a larger wedge Δ to maintain incentives.

An important example of noise in the context of banking is that of runs. Random runs lessen the connection between asset quality and survival. We know, however, that runs are not arbitrary: they are more likely to happen when asset quality is lower (Gorton, 1988; Calomiris and Gorton, 1991). From an ex-ante perspective, then, the expected risk of a run decreases when a bank chooses a safer balance sheet and this is all that matters for incentives. The fact that some good banks randomly suffer from runs does not alter this conclusion.[14]

Designation of SIFIs A separate issue is that of heterogeneity of business models. Banks, for instance, receive support from insured deposits and discount window loans that broker/dealers may not receive while insurance companies have their own risk profiles. We have already discussed how the model can deal with imperfect substitution of activities and heterogeneity in systemic risk. More generally it is conceptually straightforward to design tournaments with handicaps that depend on ex-ante heterogeneity. It is important, however, to ensure that market participants understand the actions of the government. This provides a rationale for maintaining a list of systemic firms to clarify the scope of the policy.

The main point of our paper is that tournaments can provide high-powered incentives even when the government lacks commitment. The usual limitations to the use of strong incentives are still present, as in the multitasking framework of Holmström and Milgrom (1991). Tournaments may induce banks to manipulate the measures used 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.

Rewards vs Punishments Incentives depend on the difference Δ between the "transfers" received by strong bank and weak banks. The government can increase Δ by adjusting both sides of the equation. Limited liability, we discussed in Section 3.1, puts a floor on punishment. Political constraints may put a ceiling on rewards. In general it is efficient for the government to use both rewards and punishments. In the case of mergers, it is efficient to set a low price for the failed bank and, if necessary, to subsidize the acquisition by the strong bank. With equity injections it is efficient to impose punitive terms on bad banks. In both cases it is efficient to push the value of shareholders of bad banks as low as possible, including expropriation (payment below market value). One should also emphasize that much has changed since 2008. It was difficult then to write down the value of shareholders and junior creditors without filing for Chapter 11.

[14] From an ex post perspective, punishing the weakest banks may amplify the runs they are facing. At the same time the strongest banks receive an inflow of deposits driven by a flight to safety. As our model highlights, whether the punishment remains credible depends on the substitutability between weak and strong banks, which we analyze in Section 3.2.

Today governments have resolution authority and living wills.

We have abstracted from governance conflicts within banks but these conflicts matter for the interpretation of Δ. Clawbacks and restrictions on executive compensation may be particularly effective in a world where managers do not maximize shareholders' value. The threat of nationalization and forced changes in management can also provide powerful incentives. The main difference with our theoretical model is that these punishments – unlike lowering the sale price in an acquisition – are typically not transferable to good banks. Incentives may then end up being one sided, with harsh terms imposed on bad banks while good ones go about their business.

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.

Baron, Matthew and Wei Xiong (2017), "Credit Expansion and Neglected Crash Risk*", *The Quarterly Journal of Economics*, Vol. 132, pp. 713–764.

Bengui, Julien and Javier Bianchi (2022), "Macroprudential policy with leakages", *Journal of International Economics*, p. 103659.

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.

Calomiris, W. Charles and Gary Gorton (1991), "The Origins of Banking Panics: Models, Facts, and Bank Regulation", in R. Glenn Hubbard ed. *Financial Markets and Financial Crisis*, pp. 109–174.

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, Alexi Savov, and Philipp Schnabl (2017), "The Deposits Channel of Monetary Policy*", *The Quarterly Journal of Economics*, Vol. 132, pp. 1819–1876.

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.

——— (2020), "Shadow Banking and the Four Pillars of Traditional Financial Intermediation", *The Review of Economic Studies*, Vol. 88, pp. 2622–2653.

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.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny (2015), "Neglected Risks: The Psychology of Financial Crises", *American Economic Review*, Vol. 105, pp. 310–14.

Gorton, Gary (1988), "Banking Panics and Business Cycles", *Oxford Economic Papers*, Vol. 40, pp. 751–781.

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

Hanson, Samuel G., Anil K. Kashyap, and Jeremy C. Stein (2011), "A Macroprudential Approach to Financial Regulation", *Journal of Economic Perspectives*, Vol. 25, pp. 3–28.

Holmström, Bengt (1979), "Moral Hazard and Observability", *Bell Journal of Economics*, Vol. 10, pp. 74–91.

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.

Mücke, Christian, Loriana Pelizzon, Vincenzo Pezone, and Anjan V. Thakor (2022), "The Carrot and the Stick: Bank Bailouts and the Disciplining Role of Board Appointments", working paper.

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.

Paulson, Henry M. (2010), *On the Brink: Inside the race to stop the collapse of the Global Financial System*: Business Plus.

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*.

Philippon, Thomas and Olivier Wang (2022), "Incentive Effects of Bank Mergers", working paper.

Plantin, Guillaume (2014), "Shadow Banking and Bank Capital Regulation", *The Review of Financial Studies*, Vol. 28, pp. 146–175.

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

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.

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 Main Proofs

A.1 Proof of Proposition 1

The first-best safety is the same for both banks and solves

$$x^* = \arg\max_x p_0 2f(x) + (1-p_0)(2x + \mathbb{E}[\mathcal{V}(R)|x])$$
$$= \arg\max_x p_0 2f(x) + 2(1-p_0)(1+\gamma)x$$

hence $p_0 f'(x^*) = -(1-p_0)(1+\gamma)$ which yields (6).

Given x_1, x_2 the probability that bank 1 wins the tournament is $\mathbf{P}[r_1 > r_2] = \mathbf{P}[\epsilon_1 + x_1 - x_2 \geq \epsilon_2]$. Therefore bank 1 solves

$$\max_{x_1} p_0 a f(x_1) + (1-p_0)\left(a x_1 + \left[K - a x_1 - a x_2 - \frac{\gamma}{v}\right]\int_0^{\bar{\epsilon}} G(\epsilon_1 + x_1 - x_2)\, g(\epsilon_1)\, d\epsilon_1\right)$$

where G and g are the c.d.f. and p.d.f. of ϵ, respectively. The optimality condition is

$$p_0 f'(x_1) + (1-p_0)\left[1 - \int_0^{\bar{\epsilon}} G(\epsilon_1 + x_1 - x_2)\, g(\epsilon_1)\, d\epsilon_1 + \left[K - x_1 - x_2 - \frac{\gamma}{v}\right]\int_0^{\bar{\epsilon}} g(\epsilon_1 + x_1 - x_2)\, g(\epsilon_1)\, d\epsilon_1\right]$$

Evaluating at a symmetric equilibrium $x_1 = x_2 = x$ and using the fact that ϵ_1 is uniformly distributed we get

$$p_0 f'(x) + (1-p_0)\left[1 - \frac{1}{2}\int_0^{\bar{\epsilon}} \frac{1}{\bar{\epsilon}^2}\epsilon_1 d\epsilon_1 + \left[K - 2x - \frac{\gamma}{v}\right]\frac{1}{\bar{\epsilon}}\right] = 0$$

A.2 Proof of Proposition 2

Given x_1, x_2 the probability that $r_1 \geq r_2$ is $\mathbf{P}[\epsilon_1 + x_1 - x_2 \geq \epsilon_2]$. Therefore bank 1 solves

$$\max_{x_1} p_0 a f(x_1) + (1-p_0)\left(a x_1 + \left[K - a x_1 - a x_2 - \frac{\gamma}{v}\right]\int_0^{\bar{\epsilon}} G(\epsilon_1 + x_1 - x_2)\, g(\epsilon_1)\, d\epsilon_1\right)$$

where G and g are the c.d.f. and p.d.f. of ϵ, respectively. The optimality condition is

$$p_0 f'(x_1)+(1-p_0)\left[1-\int_0^{\bar{\epsilon}} G(\epsilon_1+x_1-x_2)\,g(\epsilon_1)\,d\epsilon_1+\left[K-x_1-x_2-\frac{\gamma}{v}\right]\int_0^{\bar{\epsilon}} g(\epsilon_1+x_1-x_2)\,g(\epsilon_1)\,d\epsilon_1\right]$$

Integrating by parts we have $\int_0^{\bar{\epsilon}} G(\epsilon_1+x_1-x_2)\,g(\epsilon_1)\,d\epsilon_1=\frac{1}{2}$ therefore

$$p_0 f'(x^{\max})+(1-p_0)\left[\frac{1}{2}+\left(\frac{K-\gamma/v}{a}-2x^{\max}\right)\right]=0$$

or

$$x^{\max}=\frac{q}{f+2q}\left[\frac{1}{2}+\frac{K-\gamma/v}{a}\right]$$

where $K=2(\kappa+d)$. This shows that $x^{\max}$ is decreasing in γ and increasing in leverage d/a.

A.3 Proof of Proposition 3

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

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

hence the first best can be implemented using (13) with

$$\delta=\frac{1+\gamma}{1-\frac{1}{N}}. \tag{16}$$

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 $\Delta V \leq \epsilon$ or

$$\delta \leq \bar{\delta} = \sqrt{\frac{2e_*}{N\gamma\sigma_r^2}\eta\epsilon}. \tag{17}$$

Combining (16) and (17), we find that a sufficient condition to implement the first best is

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

Note that e_* depends on η, and the dependence can be non-monotone. However, from (12) we have that e_* converges to 0 as $\eta \to 1$ and to some positive constant $k(\gamma)$ weakly decreasing in γ (solving $V'\{k\} = \gamma$) as $\eta \to \infty$. Thus for any $\alpha \in (0,1)$ there exists $\eta_\alpha > 1$ such that for $\eta \geq \eta_\alpha$, $e_* > k(1-\alpha)$ thus we need

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

Setting α high enough, the second term dominates, which leads to Proposition 3.

A.4 Proof of Proposition 4

No mergers[...:

$$av'(e_1+m_1) = av'(e_2+m_2) = \gamma$$
$$r_1 - d + m_1 = r_2 - d + m_2 = e_*(\gamma)$$

where $e_* = v'^{-1}(\gamma/a)$ decreases with γ, increases with a. Hence total bailout is

$$M = 2e_* - e_1 - e_2$$

Ex post, using bailouts only yields value

$$2av\left(e_*\right) - \gamma\left(2e_* - e_1 - e_2\right)$$

Ex ante, we get full moral hazard because each bank solves

$$\max_x p_0 f(x) + (1-p_0)\mathbf{E}\left[r(x,\epsilon) - d + m(\dots)\right]$$
$$\max_x p_0 f(x) + (1-p_0)e_*(\gamma)$$

hence $x = 0$.

1. Forced merger at cost τ: Ex post can decide to merge bank 2 with lower return into bank 1

   $$W = 2av(e_1+e_2+M) - \tau a$$

   and then bailout M to the merged entity if needed such that

   $$2av'(e_1+e_2+M) = \gamma$$
   $$e_1+e_2+M = E_* \equiv v'^{-1}\left(\frac{\gamma}{2a}\right)$$
   $$M = E_* - e_1 - e_2$$

   Ex post value is thus

   $$2av\left(E_*\right) - \gamma\left(E_* - e_1 - e_2\right) - \tau a$$

   Comparing to the ex post value without mergers, mergers are optimal ex post when

$$2av\left(v'^{-1}\left(\frac{\gamma}{2a}\right)\right)-\gamma\left(v'^{-1}\left(\frac{\gamma}{2a}\right)-e_1-e_2\right)-\tau a > 2av\left(v'^{-1}\left(\frac{\gamma}{a}\right)\right)-\gamma\left(2v'^{-1}\left(\frac{\gamma}{a}\right)-e_1-e_2\right)$$

$$\tau < \tau^*\left(\frac{\gamma}{a}\right)$$

where

$$\tau^*\left(\frac{\gamma}{a}\right)=2\left[v\left(v'^{-1}\left(\frac{\gamma}{2a}\right)\right)-v\left(v'^{-1}\left(\frac{\gamma}{a}\right)\right)\right]-\frac{\gamma}{a}\left(v'^{-1}\left(\frac{\gamma}{2a}\right)-2v'^{-1}\left(\frac{\gamma}{a}\right)\right)$$

For instance, using the functional form $v(e)=\sqrt{e}$ (so $\alpha=1/2$) we get a very simple expression

$$\tau^*=\frac{3}{4}\frac{a}{\gamma}.$$

Note that it doesn't depend on the realization of e_1, e_2 and therefore on the ex ante decisions x. The ex ante decisions solve

$$\max_{x_1} f(x_1)+qH(x_1,x_2)E_*$$

$$x=\frac{q}{f}\phi'^{-1}\left(\frac{\gamma}{2a}\right)$$

B Micro-foundations for V and κ

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 κa_i, creditors start running, unless the equity is replenished to κa_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 κa_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 $a_i r_i - \tilde{d}_i = \kappa a_i$ that is

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

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

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

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

$$V = -\sum_i \phi\left(\kappa a_i - e_i\right).$$

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

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

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 $\rho = \rho^*$.[15] When f is linear (more generally, when decreasing returns are not at the bank level but at the aggregate level through $f(\sum k_i)$) 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 unconstrained. 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 threshold $\kappa = \frac{L(\bar{\rho})}{N}$ such that there is no externality (V does not increase with e_i) if all banks have equity $e_i \geq \kappa a_i$.

C General Pure Systemic Risk Model

This section generalizes the simple model in Section C.4 of the main text. By "pure systemic risk" we mean a value function that depends only on the *aggregate* capital surplus of the banking sector, as in Acharya et al. (2016):

$$V\{e_i\} = V\left(\sum_i (e_i - \kappa)\right) \tag{18}$$

where V is increasing and concave. For instance, the systemic expected shortfall in Acharya et al. (2016) uses the piecewise linear case $V = \min\{0, \sum_i (e_i - \kappa)\}$. 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

[15] 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.

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.

C.1 Ex Post Optimal Bailout

Define the *aggregate* return as $R \equiv \sum_i r_{i,s}$ and the aggregate gross requirement as $K \equiv \sum_i (\kappa + d_i)$. 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

$$\mathcal{M}(K-R;\gamma) \equiv \arg\max_{M\geq 0} V(R+M-K) - \Gamma(M;\gamma). \tag{19}$$

Proposition 5. *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 $V = \min(0, E-)$ 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.

C.2 First Best

With the welfare function (18), the first best solution solves

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

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

Propositions 5 and 6 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.

C.3 Moral Hazard under No Commitment and Symmetric Bailouts

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.

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

$$\sum_i m_{i,s} = \mathcal{M}\left(K - R\right) \tag{20}$$

for all possible values of R where $\mathcal{M}\left(K - R\right)$ is defined in (19). 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 $m_{i,s} = m_{j,s}$.

When all banks of ex ante identical a symmetric bailout is one where they all get the same amount of money. In a symmetric bailout satisfying the credibility constraint (20) we must have $m_{i,s} = \frac{\mathcal{M}(R)}{N}$. The best response of bank i is therefore

$$\beta_i\left(\mathbf{x}_{-i}\right) = \arg\max_{x_i \geq 0} p_0 f\left(x_i\right) + \left(1 - p_0\right)\left\{\mathbb{E}\left[r_{i,s} \mid x_i\right] + \Omega\left(x_i; \mathbf{x}_{-i}\right)\right\} \tag{21}$$

where $\mathbf{x}_{-i}$ is the vector of safety investments by all banks except bank i, and Ω is defined as

$$\Omega\left(\mathbf{x}\right) \equiv \frac{1}{N}\mathbb{E}\left[\mathcal{M}\left(K - R\right) \mid \mathbf{x}, s \neq 0\right].$$

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.[16]

Proposition 7. *All equilibria with no commitment and symmetric bailouts have the following properties:*

(i) Lack of commitment creates strategic complementarities in risk taking: $\beta_i\left(\mathbf{x}_{-i}\right)$ is increasing.

(ii) Safety is too low ($\hat{x}_i < x_i^*$) and the probability of a systemic crisis is too high: $\Phi_N\left(K \mid \hat{\mathbf{x}}\right) > \Phi_N\left(K \mid \mathbf{x}^*\right)$.

[16] 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.

(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. 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.[17]

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

C.4 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 D.

[17] 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.

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

$$\hat{x}_1 = \beta_1\left(\Delta, x_2\right) = \arg\max_{x_1} p_0 f\left(x_1\right) + \left(1-p_0\right)\left\{\mathbb{E}\left[r_{1,s} \mid x_1\right] + \Omega\left(x_1, x_2\right) + 2\Delta \times H\left(x_1, x_2\right)\right\} \tag{22}$$

where $H\left(x_1, x_2\right) = \mathbb{E}\left[H_s\left(x_1, x_2\right) | s \neq 0\right]$. 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 7. We can then state our first main proposition.

Proposition 8. *With $N = 2$, there exists a unique $\Delta^* > 0$ such that the tournament rule $\mathscr{T}$ implements the social optimum $\left(x^*, x^*, \mathcal{M}\left(K-R\right)\right)$.*

Note that Δ^* is unique in the class of mechanisms that we consider but there are other mechanisms that can implement the first best. We know from Proposition 7, however, that all of them must use some form of relative performance evaluation. Moreover, equation (22) shows that the optimal wedge Δ^* does not depend on p_0.

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

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

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

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

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

$$\hat{x}_i = \beta_i(\Delta, \mathbf{x}_{-i}) = \arg\max_{\theta} p_0 f(x_i) + (1-p_0)\left(\mathbb{E}[r_{i,s} \mid x_i] + \Omega(x_i, \mathbf{x}_{-i}) + 2\Delta \times H^{\text{med}}_N(x_i, \mathbf{x}_{-i})\right).$$

where $H^{\text{med}}_N(x_i, \mathbf{x}_{-i}) = \mathbb{E}\left[H^{\text{med}}_{s,N}(x_i, \mathbf{x}_{-i}) \,|s \neq 0\right]$. Following the same steps as for $N = 2$ we have:

Proposition 9. *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 other rules 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.

D Heterogeneous Bank Size

In the baseline model we assume banks have identical sizes $a = 1$. We now allow for different bank sizes a_i so that the equity of bank i before bailouts is $e_i = r_i a_i - d$ and denote $A = a_1 + a_2$ the total size of the banking sector.

Given the return structure of Section 2 the first best safety x^* does not depend on size.[18] Importantly, due to the credibility constraint the wedge Δ in the tournament $\mathscr{T}$ 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:

$$m_i = \begin{cases} \frac{a_i}{A}\mathcal{M}(K-R) + \Delta & \lambda_i r_{i,s} > \lambda_j r_{j,s} \\ \frac{a_i}{A}\mathcal{M}(K-R) - \Delta & \lambda_i r_{i,s} < \lambda_j r_{j,s} \end{cases} \tag{23}$$

[18] Lemma 4 in Appendix C provides more general conditions for scale independence of the first best safety.

that compares weighted returns $\lambda_i r_i$ instead of raw returns to determine the bailout allocation for some appropriate weights λ_i.

Proposition 10. *With asymmetric bank sizes $a_1 > a_2$, under the condition*

$$\frac{a_1}{a_2}\left(\frac{\frac{a_1}{A}+\gamma}{\frac{a_2}{A}+\gamma}\right) \leq 1+\frac{\bar{\epsilon} f}{q\,(1+\gamma)} \tag{24}$$

the handicapped tournament (23) with

$$\lambda_i = a_i\left(\frac{a_i}{A}+\gamma\right)$$
$$\Delta^* = \frac{\frac{1}{2}\lambda_1 \bar{\epsilon}}{1-\frac{1}{\bar{\epsilon}}\left(\frac{\lambda_1}{\lambda_2}-1\right)\frac{q}{f}(1+\gamma)}$$

implements the first best safety.

Proposition 10 is a strict generalization of Proposition 1. If $a_1 = a_2$ then $\lambda_1 = \lambda_2$ and we are back to the simple tournament case, with the same wedge Δ^* as in (10). If $a_1 > a_2$, then with a fair tournament $\lambda_1 = \lambda_2$, the prize Δ that implements $x_2 = x^*$ would be too small relative to bank 1's size, so we would have either excessive safety by small banks or insufficient safety by large banks. The handicapped tournament $\lambda_1 > \lambda_2$ is designed so that investing in safety has a higher marginal return for the large bank through a stronger effect on the probability of winning. This is a way to compensate the fact that a given dollar wedge Δ yields weaker incentives.

The left-hand side of equation (24) is increasing in the relative size a_1/a_2 hence (24) restricts the size difference a_1/a_2 to be below some upper bound which increases with the support of idiosyncratic risk captured by $\bar{\epsilon}$. Intuitively, if there is a very large bank and a very small bank $a_1 \gg a_2$ then the moral hazard is too strong and it is not possible to sufficiently motivate the large bank by pitting it against the small one, as the required wedge Δ^* would become infinite. Any handicapped tournament with positive Δ would still be a major improvement over symmetric bailouts (i.e., $\Delta = 0$ hence transfers m_i proportional to bank size) but would not implement the first best.

We start with a lemma that clarifies when we get scale independence for the first best safety, i.e., $x_i^* = x_j^*$ for all banks in spite of size differences $a_i \neq a_j$.

Lemma 4. *Let $G_\epsilon(. \mid x_i, s)$ be the distribution of $\epsilon_i = r_{i,s} - \mathbb{E}\left[r_{i,s} \mid x_i, s\right]$ 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.

Consider the case $r_{i,s} = x_i + s + \epsilon_i$. Given $\lambda = \frac{\lambda_1}{\lambda_2}$ the best response function for bank 1 is

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

while the best response function for bank 2 is

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

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

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

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

$$r_i = x_i + s + \epsilon_i. \tag{25}$$

Then

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

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 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. The next result makes this condition more explicit.

Proposition 11. *Suppose that $N = 2$, $a_1 \geq a_2$, and returns follow (25) with ϵ_i distributed over a bounded support $[0,\bar{\epsilon}]$. Then there exists*

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

such that a handicapped tournament (23) can implement the first best safety if and only if

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

Proof. We first note that if $\lambda = \frac{a_1}{a_2}$ 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

$$H_s\left(\lambda x_1 - x_2;\lambda\right) = \int_0^{\bar{\epsilon}} G_2\left(\lambda\epsilon_1 + \lambda x_1 - x_2 - (1-\lambda)s\right)g_1(\epsilon_1)d\epsilon_1$$
$$H_s'\left(\lambda x_1 - x_2;\lambda\right) = \int_0^{\bar{\epsilon}} g_2\left(\lambda\epsilon_1 + \lambda x_1 - x_2 - (1-\lambda)s\right)g_1(\epsilon_1)d\epsilon_1$$

where G_i and g_i are the c.d.f. and p.d.f. of ϵ_i, 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

$$\int_s H_s'\left(\lambda x_1 - x_2;\lambda\right)p_s ds = \int_s\left(\int_0^{\bar{\epsilon}} g_2\left(\lambda\epsilon_1 + \lambda x_1 - x_2 - (1-\lambda)s\right)g_1(\epsilon_1)d\epsilon_1\right)p_s ds$$

is zero 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. Finally, we know that the fair tournament $\lambda = 1$ implements the first best as $\frac{a_1}{a_2} \to 1$. $\square$

E Fire Sales

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.[19] Therefore the shareholder participation constraint is

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

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.

F Renegotiation-Proof Mechanisms

In Section 3.2 we studied how tournaments perform under the notion of ϵ-commitment. 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

[19] 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".

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 $r_1 > r_2$):

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

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*

$$\frac{\partial \phi}{\partial e_2}\left(\bar{e}_1, \bar{e}_1 - \frac{1+\gamma}{h}\right) \times V'\left(\phi\left(\bar{e}_1, \bar{e}_1 - \frac{1+\gamma}{h}\right) - \phi\left(\kappa\right)\right) = \gamma \tag{26}$$

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 C.4. 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

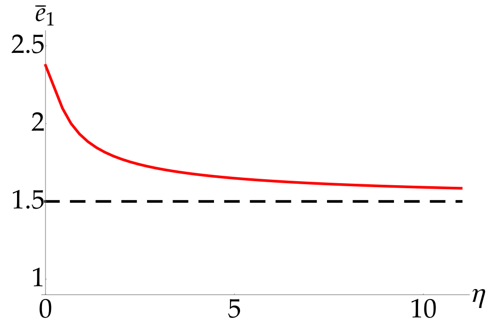

Figure 2: Renegotiation-proof prize $\bar{e}_1$ for the best bank as a function of the elasticity of substitution η. Dashed line: $\bar{e}_1$ with perfectly substitutable banks. Parameters: $V(x) = -\frac{x^2}{2}$, $\gamma = 0.5$.

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, as long as there are not other binding political constraints that put a cap on the rewards. 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 2 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 $\eta \approx 5$.

G Financial Contagion

Section 3.3 presents our results on financial contagion using an example with two banks, only one of which is systemic. In this Appendix we present the more general setup.

G.1 Earmarked Bailouts

Suppose that there are N banks and 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, unlike in the next subsection. Returns can be solved as

$$\mathbf{r} = \boldsymbol{\Lambda}\left(\mathbf{x} + \mathbf{s} + \boldsymbol{\epsilon}\right) \tag{27}$$

where $\boldsymbol{\Lambda} = \left(\mathbf{I} - \boldsymbol{\Omega}\right)^{-1}$. 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 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. The first best vector $\mathbf{x}^*$ solves

$$f'\left(x_i^*\right) = -\left(\frac{1-p_0}{p_0}\right)\lambda_i\left(1+\gamma\right). \tag{28}$$

Our baseline symmetric model is nested by setting $\boldsymbol{\Omega} = \mathbf{0}$ 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*

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

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

$$m_i = \begin{cases} \frac{\mathcal{K}}{2} + \frac{1+\gamma}{2h} - r_i & \text{if } \tilde{\lambda}_i r_i > \tilde{\lambda}_j r_j \\ \frac{\mathcal{K}}{2} - \frac{1+\gamma}{2h} - r_i & \text{if } \tilde{\lambda}_i r_i < \tilde{\lambda}_j r_j \end{cases}.$$

G.2 Contagious Bailouts

Next, we turn to 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:

$$\mathbf{r} = \mathbf{x} + \mathbf{s} + \boldsymbol{\epsilon} + \boldsymbol{\Omega}(\mathbf{r} + \mathbf{m}).$$

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

$$\mathbf{r}+\mathbf{m}=\Lambda\left(\mathbf{x}+\mathbf{m}+\mathbf{s}+\boldsymbol{\epsilon}\right)=\tilde{\mathbf{r}}+\Lambda\mathbf{m}. \tag{29}$$

The seemingly small difference relative to (27) 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 total bailout $\mathcal{M}$ 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 that bank is, the cheaper the total cost of intervening. In particular the intervention is cheaper than in the previous case of earmarker bailouts, where $\mathcal{M}=\mathcal{K}-\sum_{i=1}^N \tilde{r}_i$, if $\lambda_1>1$. Thus contagious bailouts are useful ex post because they allow the government to leverage the structure of the financial network.

However, this will backfire ex ante: when bailouts are contagious, it becomes impossible to credibly punish bank 1 and reward other banks. While this disciplines all the banks $i = 2, \ldots, N$, the countervailing force is that the most systemic bank, which should invest the most in safety in the first best allocation, is now fully insured and thus chooses the minimal safety.

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: $x_i = \tilde{x} \quad \forall i \geq 2$.*

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 $\mathcal{M} - M^* = \lambda_1 x_1^* + \sum_{i=2}^{N} \lambda_i (x_i^* - \tilde{x}) > 0$, which is increasing with λ_1.

H Other Proofs

Proof of Proposition 5. 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.

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 (19) 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

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

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 7. (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 8. 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 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

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

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

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

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

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

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:

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

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.