Bail-ins and Bail-outs: Incentives, Connectivity, and Systemic Stability

Benjamin Bernard, Agostino Capponi, and Joseph E. Stiglitz

We develop a theoretical framework to address the question of how a regulator can incentive banks to contribute to a voluntary bail-in. At the heart of the issue lies the credibility of the regulator’s threat to not bail out the financial system without the bail-in contributions of banks. We show that credible bail-in strategies exist if and only if the network hazard does not exceed a certain threshold. Incentives to join a bail-in consortium are stronger in networks where banks are more exposed to contagion, such as in sparsely connected networks, than in densely connected networks where banks know that a large part of the benefits from their bail-in contributions accrue to other banks. Our results highlight the fundamental role played by the network structure in deciding whether a rescue can be financed from sources within the banking network or whether it has to be financed by taxpayer money.

Financial institutions are linked to each other via bilateral contractual obligations and are thus exposed to counterparty risk of their obligors. If one institution is in distress, it will default on its agreements, thereby affecting the solvency of its creditors. Since the creditors are also borrowers, they may not be able to repay what they owe and default themselves—problems in one financial institution spread to others in what is called financial contagion. Large shocks can trigger a cascade of defaults with potentially devastating...
effects for the economy. The government is thus forced to intervene in some way and stop the cascade to reduce the negative externalities imposed on the economy. The extent of these cascades—the magnitude of the systemic risk—depends on the nature of the linkages, i.e., the topology of the financial system. In the 2008 crisis, it became apparent that the financial system had evolved in a way which enhanced its ability to absorb small shocks but made it more fragile in the face of a large shock. While a few studies called attention to these issues before the crisis, it was only after the crisis that the impact of the network structure on systemic risk became a major object of analysis.1

Most of the existing studies analyze the systemic risk implications of a default cascade, taking into account network topology, asset liquidation costs, and different forms of inefficiencies that arise at default. Many of these models, however, do not account for the possibility of intervention to stop the cascade. There is either no rescue of insolvent banks or the regulator and/or financial institutions intervene by following an exogenously specified protocol. The goal of our paper is to endogenize the intervention mechanism as the outcome of the strategic interaction between regulator and financial institutions.

The most common default resolution framework is the bailout, in which the government injects liquidity to help distressed banks servicing their debt. For example, during the global financial crisis, capital was injected into banks to prevent fire sale losses, including the intervention of the Bank of England and the U.S. Treasury Department’s Asset Relief Program (TARP); see also Duffie (2010) for a related discussion.2 Since the East Asia crisis, critics of bailouts have suggested bail-ins as an alternative, which are financed through voluntary contributions by the banks within the network.3 A bailed-in bank

---

1 Most notably, Allen and Gale (2000) and Greenwald and Stiglitz (2003). See also Boissay (2006), Boss et al. (2004), Castiglionesi (2007), May, Levin and Sugihara (2008), and Nier et al. (2007). One of the reasons for the limited study is the scarce availability of data on interbank linkages. An early construction of Japan’s interbank network, done before the crisis but published afterwards, is De Masi et al. (2011). With the exception of Haldane at the Bank of England, remarkably, central bankers paid little attention to the interplay of systemic risk and network topology; see Haldane (2009).

2 The Bush administration bailed out large financial institutions (AIG insurance, Bank of America and Citigroup) and government sponsored entities (Fannie Mae, Freddie Mac) at the heart of the crisis. The European Commission intervened to bail out financial institutions in Greece and Spain.

3 Duffie and Wang (2017) consider prioritization of bail-ins done contractually, rather than by a central planner. In their model, the bail-in can prioritize the counterparties of a bank, which is about to enter insolvency, that will have their debt asset positions converted into equity. Different from our paper, that
reduces payments made to creditors in exchange for equity in the reorganized company, effectively enabling the lenders to save some of their investment and the banks to stay solvent. Bail-ins alleviate the burden for taxpayers and place it on creditors of the distressed banks instead. A prominent example of a bail-in is the consortium organized by the Federal Reserve Bank of New York to rescue the hedge fund Long-Term Capital Management. As a third default resolution mechanism we consider subsidized bail-ins, where the regulator provides some liquidity assistance to incentivize the formation of bail-in consortia. Such a strategy strikes a balance between the contributions of private creditors and taxpayers. Subsidized bail-ins have been used in the recent financial crisis. The objective of our paper is to investigate the structure of default resolution plans that arise in equilibrium, when the regulator cannot credibly commit to an ex-post suboptimal resolution policy. We show how the network structure affects the regulator’s negotiation power and the banks’ incentives to participate in a rescue.

We model the provision of liquidity assistance as a sequential game between regulator and banks that consists of three stages. We consider an ex-post scenario, that is, banks have already observed the realization of (non-interbank) asset returns at the beginning of the game. In the first stage, the regulator proposes a subsidized bail-in allocation policy, specifying the quantity of debt of insolvent institutions that should be purchased by each solvent bank, as well as the additional liquidity injections that he wishes to provide to each bank (subsidies). In the second stage, each bank decides whether or not to accept

prioritization of bail-ins is viewed on the type of instrument, and not as the one based on mitigating failure contagion through a network. They show that such a prioritization has important ex-ante incentive effects on the establishment of bilateral exposures in the network.

4Long-Term Capital Portfolio collapsed in the late 1990s. On September 23, 1998, a recapitalization plan of $3.6 billion was coordinated under the supervision of the Federal Reserve Bank of New York. A total of sixteen banks, including Bankers Trust, Barclays, Chase, Credit Suisse, Deutsche Bank, Goldman Sachs, Merrill Lynch, Morgan Stanley, Salomon Smith Barney, UBS, Société Général, Paribas, Crédit Agricole, Bear Stearns, and Lehman Brothers originally agreed to participate. However, Bear Stearns and Lehman Brothers later declined to participate and their agreed-upon contributions was instead provided by the remaining fourteen banks.

5A noticeable example of a subsidized bail-in is Bear Stearns. JPMorgan Chase and the New York Federal Reserve stepped in with an emergency cash bailout in March, 2008. The provision of liquidity by the Federal Reserve was taken to avoid a potential fire sale of nearly U.S. $210 billion of Bear Stearns’ assets. The Chairman of the Fed, Ben Bernanke, defended the bail-in by stating that Bear Stearns’ bankruptcy would have affected the economy, causing a “chaotic unwinding” of investments across the U.S. markets and a further devaluation of other securities across the banking system.
the regulator’s proposal. If all banks accept, the game ends with the proposed rescue consortium and financial contagion is stopped; otherwise it moves to the third stage where the regulator is confronted with three choices: (i) purchase the debt that was supposed to be bought by the banks which rejected the proposal, (ii) purchase the entire debt, i.e., resort to a public bailout, or (iii) avoid any rescue and let the default cascade occur. After transfers are made, liabilities of banks are cleared simultaneously in the spirit of Eisenberg and Noe (2001). Unlike their paper, however, bankruptcies in our setting are costly; similarly to the papers by Rogers and Veraart (2013) and Battison et al. (2016). When there are no bankruptcy costs, the system is “conservative” and the clearing of liabilities simply reduces to a redistribution of wealth in the network. In the presence of bankruptcy costs, there are real losses that propagate through the financial system as banks cannot fully honor their liabilities. These losses are being amplified through negative feedback effects between connected defaulting banks. The size of the amplification, also referred to as the network hazard in Erol (2016), depends on the network structure.

The regulator’s option of standing idly by in the last stage is what we call the regulator’s threat of no intervention. The threat, however, may not be credible if walking away from the proposal decreases the regulator’s welfare function. We show that the credibility of the regulator’s threat is tightly linked to the topology of the financial system: the threat is credible if and only if the network hazard does not exceed a certain threshold. Our first result thus shows that the no-intervention threat is generally less credible in densely connected networks. This is because a high degree of interconnectedness between defaulting banks contribute positively to the amplification of the shock through the network. A high amplification leads to large social losses, which makes the threat non-credible and leaves a public bailout as the only possible rescue option. If, by contrast, the network hazard is small as in the case of a sparsely connected network, the regulator’s threat is credible and a subsidized bail-in may be coordinated. Our second result characterizes the welfare-maximizing subsidized bail-in that arises as the generically unique subgame perfect equilibrium (up to equivalent proposals) if the threat is credible. Our result shows
that the regulator is more likely to successfully organize a rescue consortium in a network structure that allows bail-in plans with benefits more targeted to the contributors. The intuition underlying this result is that the amplification of the shock through the network implies the existence of a network multiplier: the gains from a rescue for each dollar injected into the financial system is generally greater than one dollar. These gains, however, accrue to the entire financial sector. An individual bank is willing to contribute if and only if the share of these gains accrued to the contributing bank itself is also greater than one (conditional on every other bank’s contribution decision).

By analyzing the ring and complete network as representative structures of sparsely and densely connected networks, respectively, we show that the ring network allows for a better targeted bail-in and is thus welfare enhancing. The intuition behind this result is that a complete network, in which each bank is equally liable to every other bank in the system, is susceptible to free-riding. As all banks are equally exposed to contagion, they know that a large part of the gains from their contributions will accrue to other banks. Hence, the same exact force that creates an absorption mechanism for losses in a complete network (Allen and Gale (2000) and Acemoglu, Ozdaglar and Tahbaz-Salehi (2015)) leads to free-riding when the banks are faced with the decision of how much to contribute to a bail-in. On the other hand, in a ring network each bank has exactly one creditor, hence banks can be ranked based on their exposure to contagion. Those which are the most exposed are precisely the same banks which would get most of the benefits from their social contributions to a bail-in, and thus have stronger incentives to join a bail-in. The impacts of these forces on equilibrium welfare losses are illustrated in Figure 1.

Our analysis indicates that the regulator’s threat of inaction is credible only if the amplification of the shock is smaller than the social losses due to asset liquidation. This result contributes to explain some of the decisions made by

---

6Empirical research, see for instance Craig and Von Peter (2014), has shown that the network of interbank obligations exhibits a core-periphery structure. The core banks are the large and systemically important financial institutions, while periphery banks are small institutions which primarily trade with core banks but do not have contractual obligations to each other. Our comparison between the ring and the complete network is relevant for the subnetwork induced by the core banks. Since periphery banks do not trade with each other, a fundamental default of a periphery bank can be modelled as an exogenous shock hitting the subnetwork of core banks.
Figure 1: The figure compares equilibrium welfare losses in the complete (blue) and the ring network (red) in the presence (solid lines) and absence (dashed lines) of intervention. When the no-intervention losses exceed the costs of a public bailout (black dashed line), the government’s threat to not intervene is not credible and a public bailout is the only possible equilibrium rescue. Banks’ contributions are larger in the ring network.

the sovereign authorities during distress scenarios. For instance, a private bail-in was coordinated to rescue the Long-Term Capital Management hedge fund in 1998. In contrast, the government of the United States rescued Citigroup through a public bailout in November 2008, preventing the bankruptcy of the largest bank in the world at the time. Compared to the period when Long-Term Capital Management was rescued, the capitalization of the entire financial system was much lower at the time when Citigroup entered into distress, due to the many default events which had already occurred prior to Citigroup’s bailout. Because the amplification of a shock is higher in a lowly capitalized network, the credibility of the no-intervention threat is lower.

The remainder of the paper is organized as follows. In Section 1, we provide a review of the existing literature before developing the details of our model in Section 2. We characterize the optimal bail-in plan and the equilibrium outcome for a fixed network in Section 3. We analyze the impact of the network topology on the credibility of the regulator’s threat in Section 4. Section 5 illustrates and interprets our main results with examples and Section 6 provides concluding remarks. In Appendix A, we provide a generalization of the results of Section 3. All proofs are delegated to appendices B–D.

7The figure displays our results in the stylized case of a continuum of banks to highlight the key differences between welfare losses in a ring and a complete network. In our model, the financial network consists of a finite number of banks, which leads to additional discontinuities in welfare losses. We refer to Section 5 for a numerical example consisting of a finite number of banks.

8We recall the seven credit events occurred in the month of September 2008, involving Fannie Mae, Freddie Mac, Lehman Brothers, Washington Mutual, Landsbanki, Glitnir and Kaupthing.
1 Literature Review

Our paper is related to a vast branch of literature on financial contagion in interbank networks. Pioneering works include Allen and Gale (2000) and Eisenberg and Noe (2001), and further developments were made in more recent years by Acemoglu, Ozdaglar and Tahbaz-Salehi (2015), Elliott, Golub and Jackson (2014), Gai, Haldane and Kapadia (2011), Glasserman and Young (2015), and Capponi, Chen and Yao (2016). We refer to Glasserman and Young (2016) for a thorough survey on financial contagion. These works study how an initial shock is amplified through the interbank network based on the network topology, the impact of inefficient liquidation of non-interbank assets, the trade-off between diversification and integration on the levels of interbank exposures, and how greater complexity of the network may increase its fragility, making it vulnerable to the spread of contagion. Different from our study, in these models agents execute exogenously specified contractual agreements but do not take any strategic action to resolve distress in the network. A different branch of the literature has focused on the coordinating role of the central bank in stopping financial contagion through provision of liquidity to solvent but illiquid institutions (Freixas, Parigi and Rochet (2000) and Gorton (2010)). These studies, however, do not consider resolution strategies featuring the involvement of the private sector, such as the bail-ins considered in our paper. The presence of this default resolution option questions the credibility of the no-intervention commitment by the government, which we show to be tightly linked to the network structure and the size of the shocks hitting the banking system.

In our model, the regulator determines the optimal rescue plan after the banks have already decided on the amount of risk undertaken, and shocks to...
outside assets have occurred. We do not consider here the moral hazard problem of banks taking excessive risks, knowing that they will be rescued if the market moves against them. Moral hazard problems arising in this context have been thoroughly investigated in the literature, but the proposed models do not account for the structure of the interbank network (the exception is a recent study by Erol (2016), which develops an endogenous network formation model in which the government can intervene to stop contagion). These studies include Gale and Freixas (2002), who argue that a bailout is optimal ex post, but ex ante it should be limited to control moral hazard; Acharya and Yorulmazer (2007), who show that banks may find it optimal to invest in highly correlated assets in anticipation of a bailout triggered by the occurrence of many simultaneous failures; Farhi and Tirole (2012), who support Acharya and Yorulmazer (2007)’s findings by showing that safety nets can provide perverse incentives and induce correlated behavior that increases systemic risk; Chari and Kehoe (2016), who show that if the regulator cannot commit to avoid bailouts ex post, then banks may overborrow ex ante; and Keister (2016), who finds that prohibiting bailouts may lead intermediaries to invest into too liquid assets which lower aggregate welfare. In contrast with all these studies, banks do not strategically decide on interbank or outside asset investments in our model, but the channel of contagion comes primarily from the propagation of shocks through the exogenous network of liabilities.

Related to our model is the study by Rogers and Veraart (2013), who analyze situations in which banks can stop the insolvency from spreading by stabilizing the financial system through mergers. The question of whether such a merger is incentive compatible for the shareholders of an individual bank is not addressed, however, and the government does not take an active role in their paper. Differently from Rogers and Veraart (2013), our model thus allows us to address questions such as the credibility of the regulator’s actions and the free-riding problem that arises because the stability of the financial system is to the benefit of every participant.
2 Model

We consider an interbank network with simultaneous clearing in the spirit of Eisenberg and Noe (2001). Banks $i = 1, \ldots, n$ are connected through interbank liabilities $L = (L^{ij})_{i,j=1,\ldots,n}$, where $L^{ij}$ denotes the liability of bank $j$ to bank $i$. We denote by $L^j := \sum_{i=1}^n L^{ij}$ the total liability of bank $j$ to other banks in the network. Define the relative liability matrix $\pi = (\pi^{ij})$ by setting $\pi^{ij} = L^{ij}/L^j$ if $L^j \neq 0$ and $\pi^{ij} = 0$ otherwise. Our framework can accommodate lending from the private sector by adding a “sink node” $n + 1$ that has only interbank assets but no interbank liabilities. Banks have investments in outside assets with values $e = (e^1, \ldots, e^n)$, cash holdings $c_h = (c^1_h, \ldots, c^n_h)$ and financial commitments $c_f = (c^1_f, \ldots, c^n_f)$ with a higher seniority than the interbank liabilities. These commitments include wages and other operating expenses. If a bank $i$ is not able to meet its liabilities $L^i + c^i_f$ out of current income, it will liquidate a part $\ell^i \in [0, e^i]$ of its outside investments, but will recover only a fraction $\alpha \in (0, 1]$ of its value.\(^9\)\(^10\) If a bank $i$ cannot meet its liabilities even after liquidating all of its outside assets $e^i$, it will default. As in Rogers and Veraart (2013) and Battison et al. (2016), the default of a bank is costly and only a fraction $\beta \in (0, 1]$ of the defaulting bank’s value is paid to the creditors. We use $A = \pi L$ to denote the vector of book values of interbank assets with $A^i = (\pi L)^i = \sum_{j=1}^n \pi^{ij} L^j$. Because the operating expenses $c_f$ have higher seniority than the interbank liabilities, the model depends on $c_h$ and $c_f$ only through the net cash balance $c = c_h - c_f$. For the sake of brevity, denote by $V_0 = c + e + A - L$ the book value of bank $i$’s equity.

We denote by $(L, \pi, e, c)$ the financial system after risks have been taken and after an exogenous shock has hit the system. The shock may lower the value of banks’ outside assets or the net cash holdings of the banks. This may result in a negative net cash balance if, for example, a bank intended to use the returns from an investment to cover the operating expenses, but the returns turned out to be lower than expected. We refer to the defaults that occur as an

\(^9\)In reality, recovery rates are asset-specific and some assets may directly be transferred to the creditors of defaulting institutions without liquidation. The parameter $\alpha$ is to be understood as an average recovery rate across all assets. It is equal to 1 if all assets are transferred to the creditors.
immediate consequence of the shock as the fundamental defaults and denote their index set by \( \mathcal{F} := \{ i \mid L^i > c^i + \alpha e^i + A^i \} \). These are banks which cannot meet their obligations even if every other bank repays its liabilities in full. Because fundamentally defaulting banks are able to only partially repay their creditors, their defaults may lead to additional defaults in the system, resulting in a default cascade. If, however, banks in \( \mathcal{F} \) receive a liquidity injection so that they can meet their obligations, the financial system is stabilized. In this section, we first characterize the outcome of a default cascade and then discuss in detail the different types of rescue under consideration.

### 2.1 Default cascade

A defaulting bank will recall its assets and repay its creditors according to their seniority. Depositors are the most senior creditors, hence they are given priority over lenders from the interbank network and the private sector, to whom we refer as junior creditors. Creditors with the same seniority are repaid proportionally to their claim sizes. How much a bank is able to recall from its interbank assets depends on the solvency of the other banks in the system.

A clearing payment vector is a set of repayments, simultaneously executed by all banks, for which every solvent bank repays its liabilities in full and every insolvent bank precisely repays its total value after liquidation\(^{11}\).

**Definition 2.1.** A clearing payment vector \( p = (p^1, \ldots, p^n) \) for a financial system \((L, \pi, e, c)\) is a fixed point of

\[
p^i = \begin{cases} 
L^i & \text{if } c^i + \alpha e^i + \sum_{j=1}^n \pi^{ij} p^j \geq L^i, \\
\beta \left( c^i + \alpha e^i + \sum_{j=1}^n \pi^{ij} p^j \right)^+ & \text{otherwise}\end{cases}
\]

For a clearing payment vector \( p \), we denote by \( \mathcal{D}(p) := \{ i \mid p^i < L^i \} \) and \( \mathcal{S}(p) := \{ i \mid p^i = L^i \} \) the set of defaulting and solvent banks, respectively. If

\(^{11}\)In practice, liabilities may be cleared sequentially rather than simultaneously and the order of clearing may impact the outcome. This method of simultaneous clearing is standard in the literature and may represent the fact that clearing of liabilities occurs on a much smaller time scale than the formation of rescue consortia.

\(^{12}\)We use \( x^+ = \max(x, 0) \) and \( x^- = \max(-x, 0) \) to denote the positive and negative part of \( x \), respectively.
clearing payments are given by $p$, each bank $i$ liquidates a nominal amount equal to

$$
\ell^i(p) = \min \left( \frac{1}{\alpha} \left( L^i - c^i - \sum_{j=1}^{n} \pi^{ij} p^j \right)^+, e^i \right).
$$

That is, each solvent bank liquidates just enough to remain solvent and each defaulting bank liquidates the outside assets in their entirety. A defaulting bank pays its entire value to its creditors. If the payment $p^i$ is positive, it is divided pro-rata among bank $i$’s junior creditors and the senior creditors (e.g., the depositors) are paid in full. If $p^i = 0$, the junior creditors do not receive anything and the senior creditors suffer a loss of

$$
\delta^i(p) := \left( c^i + \alpha e^i + \sum_{j=1}^{n} \pi^{ij} p^j \right)^-.
$$

Our notion of clearing payment vectors extends the corresponding notion in Rogers and Veraart (2013), by allowing banks to partially liquidate their outside assets, and the corresponding notion in Acemoglu, Ozdaglar and Tahbaz-Salehi (2015) by incorporating bankruptcy costs. The value of bank $i$’s equity after liabilities are cleared with clearing payment vector $p$ is equal to

$$
V^i(p) := (\pi p + c + e - (1 - \alpha)\ell(p) - p^i) 1_{\{p^i = L^i\}}.
$$

We define the welfare losses $w_\lambda(p)$ as the weighted sum of the losses due to default costs, that is,

$$
w_\lambda(p) = (1 - \alpha) \sum_{i=1}^{n} \ell^i(p) + \frac{1 - \beta}{\beta} \sum_{i \in D(p)} p^i + \lambda \sum_{i \in D(p)} \delta^i(p).
$$

The first two terms in (1) are deadweight losses due to inefficient asset liquidation and bankruptcy costs: If bank $i$ liquidates a positive amount $\ell^i(p)$, only $\alpha \ell^i(p)$ is recovered and $(1 - \alpha)\ell^i(p)$ is lost. A defaulting bank $i$ repays a fraction $\beta$ of its assets $p^i/\beta$ before bankruptcy, hence $(1 - \beta)p^i/\beta$ is lost. The last term in (1) are the losses borne by senior creditors, weighted by a constant.
\( \lambda \geq 0 \). The weight \( \lambda \) captures the importance that the regulator assigns to the senior creditor’s losses relative to the deadweight losses from asset liquidation and bankruptcy proceedings. The regulator’s goal is to minimize welfare losses and the parameter \( \lambda \) captures his priorities in doing so. A regulator with \( \lambda = 0 \) views the senior creditors’ losses merely as transfers of wealth and not as losses to the economy. A higher value of \( \lambda \) indicates a higher priority to the welfare of the economy exclusive of the banking sector. The coefficient \( \lambda \) may also be interpreted as a measure of political pressure on the regulator since a haircut on deposits leads to unhappy voters.

We obtain the following existence result for clearing payment vectors in analogy with Rogers and Veraart (2013).

**Lemma 2.1.** For any financial system \( (L, \pi, c, e) \), there exist a greatest and a lowest clearing payment vector \( \bar{p} \) and \( p \), respectively, with \( \bar{p}^i \geq p^i \geq p^i \) for any clearing payment vector \( p \) and any bank \( i \). Moreover, \( \bar{p} \) is Pareto dominant, i.e., \( V^i(\bar{p}) \geq V^i(p) \) and \( w_\lambda(\bar{p}) \leq w_\lambda(p) \) for any \( \lambda > 0 \).

As is standard in the literature, liabilities are cleared with the Pareto-dominant clearing payment vector \( \bar{p} \). For two sets \( S, I \subseteq \{1, \ldots, n\} \), denote by \( \pi_{S,I} \) the submatrix of \( \pi \) given by the rows in \( S \) and the columns in \( I \). Similarly, given a vector \( x \), we use \( x_S \) to denote the subvector of \( x \) with entries in \( S \). We will often use the 1-norm to denote the sum over the absolute entries of a vector, i.e.,

\[
\|x\|_1 = \sum_{i=1}^{n} |x^i|, \quad \|x_S\|_1 = \sum_{i \in S} |x^i|.
\]

### 2.2 Coordination of rescues

**Definition 2.2.** A **subsidized bail-in** \((b, s)\) consists of a bail-in allocation \( b = (b^1, \ldots, b^n) \) and a vector of subsidies \( s = (s^1, \ldots, s^n) \) that specifies for each bank \( i = 1, \ldots, n \) the amount \( b^i \) of debt that bank \( i \) purchases from fundamentally defaulting banks and the size \( s^i \) of the subsidy bank \( i \) receives. The government’s contribution to the bail-in is \( \|s\|_1 - \|b\|_1 \).
Remark 2.1. Subsidized bail-ins contain public bailouts and privately backed bail-ins as special cases. A public bailout is a subsidized bail-in, in which the banks’ contributions are equal to 0, that is, \( b = 0 \). In a private bail-in, the government contributions are 0, i.e., \( \|b\|_1 = \|s\|_1 \).

In addition to contributing to a bail-in financially, the regulator also serves to coordinate among different bail-ins. Specifically, the regulator may propose a bail-in strategy, but cannot force banks to participate.\(^{13}\) The organization of a rescue thus takes the following stages.

1. The regulator proposes a subsidized bail-in \((b, s)\).

2. Each bank \( i \in \mathcal{F}^c \) chooses a binary action \( a^i \in \{0, 1\} \), indicating whether or not it accepts the proposed bail-in.\(^{14}\) Let \( \mathcal{R} := \{ i \in \mathcal{F}^c \mid a^i = 0 \} \) denote the set of banks which reject the proposal.

3. The regulator has the following three options:

   (a) \( a^0 = \mathcal{R} \): Proceed with the planned subsidies \( s \) without the contributions of banks in \( \mathcal{R} \). Bank \( i \)'s net cash balance in the rescue is \( c^i(b, s, a) = c^i + s^i - b^i1_{\{a^i=1\}} \). Let \( \bar{p}(b, s, a) \) denote the Pareto-dominant clearing payment vector in the financial system \((L, \pi, e(b, s, a), e)\). The value of bank \( i \) after the rescue is

\[
V^i(b, s, a) := V^i(\bar{p}(b, s, a)).
\]

Welfare losses are equal to

\[
w(\lambda)(b, s, a) = w(\lambda)(\bar{p}(b, s, a)) + \lambda \sum_{i=1}^{n} (s^i - b^i1_{\{a^i=1\}}). \quad (2)
\]

\(^{13}\)Duffie and Wang (2017) consider bail-in strategies which are done contractually, rather than by a central planner. In their model, prioritization of bail-ins is viewed as based on the type of the instrument, and not as one based on mitigating failure contagion through the network. Under strong axioms and assumptions on bilateral bargaining conditions, they show that the efficient choice of bail-in arrangements is made voluntarily.

\(^{14}\)For a set of banks \( \mathcal{X} \), we denote by \( \mathcal{X}^c := \{1, \ldots, N\} \setminus \mathcal{X} \) the complement with respect to \( \{1, \ldots, N\} \).
(b) $a^0 = P$: Resort to a public bailout $(0, \bar{s})$ of the regulator’s choice. With the subsidies, the net cash balance of bank $i$ after the intervention is equal to $c^i(\bar{s}) = c^i + \bar{s}^i$. We denote by $\bar{p}(\bar{s})$ the greatest clearing payment vector in the financial system $(L, \pi, c(\bar{s}), e)$. The value of the banks and the welfare losses are identical to (a).

(c) $a^0 = N$: Abandon the rescue, which results in a default cascade as in Section 2.1. We denote the welfare losses in a default cascade by $w_N := \lambda(\bar{p}(0, 0, 0))$ for the sake of brevity. Observe that a default cascade technically coincides with a bailout with 0 subsidies.

**Definition 2.3.** A bail-in $(b, s)$ is feasible if it does not require any contributions of fundamentally defaulting banks and each non-fundamentally defaulting bank $i$ can afford to contribute $b^i$, given subsidies $s$. More precisely, $b^i = 0$ for any $i \in F$ and $L^i + b^i \leq c^i + s^i + \alpha c^i + (\pi \bar{p}(b, s, 1))^i$ for $i \not\in F$. A bail-in is thus feasible if it can be accepted by all banks.

The goal of the next section is to characterize all subgame perfect equilibria of this game when the regulator is restricted to propose feasible bail-ins. Subgame perfection eliminates the non-credible threat of the regulator to abandon the rescue in the third stage when, in fact, he prefers a public bailout over a default cascade. This makes it impossible to incentivize bank $i$ to participate in a bail-in if welfare losses are lowest in a bailout that protects bank $i$. Indeed, knowing that the regulator will inevitably resort to a public bailout after a rejection of his proposal, bank $i$ has no punishments to fear for not cooperating.

### 3 Characterization of equilibria

In this section, we characterize the set of equilibria via backward induction. To get across the intuition behind our results as clearly as possible, we focus on all-or-nothing interventions, where the regulator considers only rescues in which every bank of the system is saved. The reason why pursue this approach is that the mathematical complexity is reduced without sacrificing much of the economic intuition behind the results. With our approach, the network’s structure impacts the coordination of bail-ins only through the credibility of the
regulator’s threat, which stands at the heart of this analysis. The payoffs in a rescue, however, decouple from the network structure because in a complete rescue, every bank recovers the book value $A^i$ of its interbank assets. If general rescues are allowed, the regulator simply performs an additional minimization in the first stage, where he considers rescues of all possible subsets of banks. However, for each subset considered, the formation of bail-ins works analogously as in the all-or-nothing approach. The results for general rescues are contained in Appendix A.

3.1 Public bailout

In a complete bailout, the regulator provides a subsidy to every distressed bank. For a fundamentally defaulting bank $i$, there are two potentially sensible choices for the size of the subsidy $s^i$. The regulator can either give a subsidy that is just large enough so that bank $i$ does not have to liquidate any of its investments or he can give a subsidy that is just large enough to prevent bank $i$ from defaulting, but that requires bank $i$ to liquidate its projects. Which of the two choices minimizes the welfare losses depends on the liquidation costs of the asset and on the value of taxpayer money.

In a complete rescue, each bank recovers its interbank assets in full. The nominal amount that any bank $i$ has to liquidate to recover cash amount $x$ thus decouples from the network structure. The amount bank $i$ has to liquidate is

$$\ell_R^i(x) = \min\left(\frac{1}{\alpha}(L^i + x - c^i - A^i)^+ + e^i, e^i\right).$$

(3)

Lemma 3.1. The welfare-maximizing complete bailout $s_P$ is given by

$$s_P^i = \begin{cases} 
(L^i - c^i - A^i)^+ & \text{if } \lambda \alpha < 1 - \alpha, \\
(L^i - c^i - \alpha e^i - A^i)^+ & \text{if } \lambda \alpha \geq 1 - \alpha.
\end{cases}$$
The equity value of each bank $i$ in $s_P$ is

$$V^i_P := \begin{cases} 
V^i_0 + s^i_P & \text{if } \lambda \alpha < 1 - \alpha, \\
V^i_0 - (1 - \alpha)\ell^i_R(0) & \text{if } \lambda \alpha \geq 1 - \alpha,
\end{cases} \quad (4)$$

We denote the welfare losses by $w_P = w_\lambda(0, s_P, 1)$. Welfare losses can be expressed as $w_P = \lambda B + \min(\lambda \alpha, 1 - \alpha)\|\ell_R(0)\|_1$, the weighted sum of aggregate shortfall $B = \sum_{i \in F}(L^i - c^i - \alpha e^i - A^i)^+$ of fundamentally defaulting banks and an additional term $\min(\lambda \alpha, 1 - \alpha)\|\ell_R(0)\|_1$ due to subsidies or liquidation.

The interpretation behind (4) is the following. Because everybody is rescued, each bank essentially recovers its full value $V^i_0$ up to differences from subsidies and/or liquidation. If taxpayer money is cheap, i.e., $\lambda \alpha < 1 - \alpha$, then the regulator covers the entire shortfall of bank $i$ because the cost $\lambda$ of a taxpayer dollar is cheaper than the deadweight losses $(1 - \alpha)/\alpha$ per dollar of recovered asset in liquidation. If, however, taxpayer money is expensive, i.e., $\lambda \alpha \geq 1 - \alpha$, the regulator provides a subsidy that is only large enough to cover the shortfall of bank $i$ after $i$ liquidates its assets in full. Bank $i$ thus has to liquidate the amount $\ell^i_R(0)$, which is equal to $e^i$ for fundamentally defaulting banks and equal to $\frac{1}{\alpha}(L^i - c^i - A^i)^+$ otherwise.

**Corollary 3.2.** Let $(b, s)$ be the proposed bail-in with response vector $a$ of the banks. In the last stage, the regulator chooses to implement the bail-in if and only if $w_\lambda(b, s, a) < \min(w_N, w_P)$.

### 3.2 Equilibrium response of banks

For any feasible bail-in proposal, there may be many equilibrium responses. A crucial separation among equilibria is whether or not the regulator decides to go ahead with the bail-in in the last stage. Suppose that the regulator proposes a bail-in, to which at least 5 banks need to agree to in order for the regulator to prefer it over the alternatives of a default cascade or public bailout. Then any response with at least 3 banks accepting the proposal is
trivially an equilibrium because a deviation of a single bank is not going to change the outcome.

More interesting is the multiplicity among equilibria where the regulator proceeds with the bail-in. There may be several coalitions of banks, with which the regulator is happy to coordinate an intervention. While any two such equilibria may not be comparable, the regulator can preempt the coordination problem by altering his proposed bail-in so that it is incentive compatible to accept the proposal for precisely one coalition. For such a bail-in proposal, the unique accepting equilibrium Pareto-dominates all rejecting equilibria.

**Definition 3.1.**

1. An equilibrium \((b, s, a)\) is called an **accepting equilibrium** if \(a^0 = R\) and a **rejecting equilibrium** otherwise. We use the same terminology for equilibria \(a\) of the continuation game after a proposal \((b, s)\) has been made.

2. Two equilibria \((b, s, a)\) and \((\tilde{b}, \tilde{s}, \tilde{a})\) are **equivalent** if they ask the same net contribution \(b^i - s^i = \tilde{b}^i - \tilde{s}^i\) of each bank \(i = 1, \ldots, n\) and a bank’s response in \(a\) and \(\tilde{a}\) differs only if its proposed net contribution is 0.

**Lemma 3.3.** Let \((b, s)\) be a feasible bail-in proposal. In an accepting equilibrium \(a\), bank \(i\) with \(b^i > 0\) accepts if and only if the following conditions hold:

1. \(w\lambda(b, s, (0, a^{-i})) \geq \min(w_N, w_P)\)

2. \(b^i - s^i \leq \begin{cases} \sum_{j=1}^{n} \pi_{ij}(L^j - p^j_N) & \text{if } w_N \leq w_P, \\ 0 & \text{if } w_P < w_N. \end{cases}\)

Let us break down the intuition behind this result. If \(w_N \leq w_P\), then the regulator prefers a default cascade over a bailout. The first condition thus states that there is no possibility for free-riding: if bank \(i\) were to reject the proposed bail-in, the regulator is not going to make up for \(i\)'s contribution and lets a default cascade occur instead. The second condition states how much bank \(i\) is willing to contribute to prevent a default cascade. Bank \(i\) is willing to \(^{15}\)We use the standard notation in game theory, denoting by \((0, a^{-i})\) the action profile, where bank \(i\) rejects the proposal and each other player’s action (including the regulator’s) is the same as in action profile \(a\).
make a net contribution to the bail-in up to the amount the bank would lose in a default cascade. If $w_N > w_P$, the two conditions state that the regulator’s threat to not bail out the banks is not credible. Because a rejection of bank $i$ leads to a bailout by Condition 1, bank $i$ accepts only bail-ins that are a net subsidy by Condition 2. If $w_N > w_P$, the regulator can thus do no better than resorting to a public bailout.

**Lemma 3.4.** Let $(b, s)$ be a proposed bail-in plan with accepting equilibrium responses $\{a_1, \ldots, a_m\}$. For any $a_k, k = 1, \ldots, m$, there exists a proposal $(\tilde{b}, \tilde{s})$, to which $a_k$ is the unique accepting equilibrium response (up to equivalence). Moreover, $a_k$ is the Pareto-dominant equilibrium response to $(\tilde{b}, \tilde{s})$.

The significance of Lemma 3.4 is that the regulator has the power to select among accepting equilibria by modifying his/her proposal. Because the selected equilibrium is the unique accepting equilibrium for the modified proposal, the regulator can ensure that there are no coordination problems between banks.

### 3.3 Optimal proposal of the regulator

Because the regulator can anticipate the banks’ response in equilibrium, it is unnecessary for him/her to propose a bail-in that is rejected by any bank. He will thus propose the bail-in that minimizes welfare losses among all bail-ins that satisfy the conditions of Lemma 3.3. This also justifies modeling the negotiation between regulator and the banks as a single interaction. In reality, there might be several rounds of negotiation.\(^\text{16}\) If some bank rejects the regulator’s proposal, the regulator might revise it to convince other banks to join the consortium. In a game-theoretic model with complete information, the entire negotiation process is collapsed into a single stage by proposing only bail-in plans that are incentive compatible for all members of the consortium. Moreover, by Lemma 3.4, the regulator can ensure that the accepting equilibrium is unique and Pareto dominant.

\(^{16}\)When the bail-in consortium for the rescue of Long Term Capital Management was coordinated in the late 90s, Bear Stearns and Lehman Brothers rejected the bail-in proposed by the Federal Reserve Bank of New York. Their share of the bail-in was then redistributed among the remaining 14 banks.
If $w_P < w_N$, the regulator’s threat to not intervene fails to be credible. Lemma 3.3 implies that in this case, the regulator cannot motivate any bank to participate. It follows that the regulator’s optimal proposal is the optimal bailout of Lemma 3.1. If $w_P \geq w_N$, the no-intervention threat is credible and banks can be motivated to participate. Let $(b, s)$ be a feasible bail-in proposal with a unique accepting equilibrium $a = (R, 1, \ldots, 1)$. Bank $i$’s contribution to the bail-in $(b, s)$ reduces welfare losses by $\lambda b^i - (1 - \alpha)\left(\ell^i_R(b^i - s^i) - \ell^i_R(-s^i)\right)$. It lowers the regulator’s contribution by $b^i$, but may create liquidation losses for the bank to recover the contributed amount. The “no free-riding” condition of Lemma 3.3 implies that the welfare losses after the proposal’s rejection by a single bank $i$ are bounded from below by $w_N$. The welfare losses in a bail-in $(b, s)$ thus admit the lower bound

$$w(\lambda, b, s, a) \geq w_N - \min_i \left(\lambda b^i - (1 - \alpha)\left(\ell^i_R(b^i - s^i) - \ell^i_R(-s^i)\right)\right). \quad (5)$$

The regulator will thus strive to include banks in the bail-in which offer a high contribution to the rescue consortium and generate low deadweight losses when they liquidate their outside assets to retrieve the contributed amount. Which choice of $b$ is optimal for the regulator depends on the values of $\lambda$ and $\alpha$: if taxpayer money is expensive relative to liquidation costs ($\lambda \alpha \geq 1 - \alpha$), the regulator prefers that banks liquidate their outside assets to buy up a larger amount of debt, whereas if taxpayer money is cheap ($\lambda \alpha < 1 - \alpha$), the regulator prefers to buy more debt himself so as to avoid the liquidation of banks’ outside assets. We are now ready to state the main result for all-or-nothing rescues, which characterizes the optimal proposal of the regulator and the equilibrium welfare losses in any accepting equilibrium.

**Theorem 3.5.** Let $i_1, \ldots, i_{|F|}$ be a non-increasing ordering of banks according to $\nu^i := \lambda \eta^i - (1 - \alpha)\left(\ell^i_R(\eta^i - s^i_P) - \ell^i_R(-s^i_P)\right)$, where

$$\eta^i := \begin{cases} 
\min \left(\sum_{j=1}^n \pi^{ij}(L^j - p^j_N), (c^i + \alpha e^i + A^i - L^i)^+\right) & \text{if } \lambda \alpha \geq 1 - \alpha, \\
\min \left(\sum_{j=1}^n \pi^{ij}(L^j - p^j_N), (c^i + A^i - L^i)^+\right) & \text{if } \lambda \alpha < 1 - \alpha.
\end{cases}$$
Let \( m := \min \left( k \mid w_P - \sum_{j=1}^{k} \nu^{ij} < w_N \right) \). If \( w_P < w_N \), then the unique equilibrium outcome is the public bailout \( s_P \) by the regulator. If \( w_N \leq w_P \), then all accepting equilibria Pareto dominate all rejecting equilibria and the equilibrium welfare losses in any accepting equilibrium are equal to

\[
w_* = \min \left( w_P - \sum_{j=1}^{m} \nu^{ij}, w_N - \nu^{i_{m+1}} \right).
\]

Moreover, if \( w_* = w_P - \sum_{j=1}^{m} \nu^{ij} \), there exists a unique accepting equilibrium with proposal \((b_*, s_*)\), where banks \( j = i_1, \ldots, i_m \) each contribute \( \nu^j \) and subsidies are given by \( s_* = s_P \). If \( w_* = w_N - \nu^{i_{m+1}} \), a proposal \((b_*, s_*)\) is part of an accepting equilibrium if and only if each bank \( j = i_1, \ldots, i_{m+1} \) contributes \( \nu^j \) and subsidies satisfy \( s_* \geq s_P \) subject to

1. \( b_j^* - s_j^* \geq \nu^{i_{m+1}} \) for any \( j = i_1, \ldots, i_{m+1} \), and
2. \( \|s_*\|_1 - \|s_P\|_1 = \frac{1}{\lambda} \left( w_P - w_N - \sum_{j=1}^{m} \nu^{ij} + \nu^{i_{m+1}} \right) \).

Theorem 3.5 is best explained in words as follows. If the no-intervention threat of the regulator fails to be credible, then the unique equilibrium outcome is the public bailout characterized in Lemma 3.1. If the no-intervention threat is credible, the organization of a bail-in is consistent with equilibrium behavior. The subsidies that each bank receives in the optimal bail-in are the same as in the optimal bailout since Theorem 3.5 considers all-or-nothing rescues only.

The quantity \( \eta^i \) is the largest contribution by bank \( i \) that the regulator is willing to request and the bank is willing to pay. It is essentially given by the losses of bank \( i \) in a default cascade up to bank \( i \)'s budget constraint with liquidation \( (\lambda \alpha \geq 1 - \alpha) \) and without liquidation \( (\lambda \alpha < 1 - \alpha) \). The quantity \( \nu^j \) is the welfare impact of a contribution \( \eta^i \) by bank \( i \) to a bail-in: the regulator can save \( \eta^i \) taxpayer dollars, which improves welfare by \( \lambda \eta^i \). However, there may be a cost associated for the bank to recover the quantity \( \eta^i \). Bank \( i \) has to liquidate a nominal amount \( \ell^i_R(\eta^i - s^i_P) \) to be able to afford \( \eta^i \), which generates deadweight losses. Note that the liquidated amount could be 0.
A banks’ contributions to the rescue of insolvent banks benefits other financial institutions in the system through the stability of the financial network. As such, the coordination of bail-ins has an inherent free-riding problem, where each bank prefers to let others contribute in its stead. By adding banks to the rescue consortium in the decreasing order \(i_1, i_2, \ldots\), the regulator asks for contributions of banks first, which get the largest benefit from the rescue. These are banks with the lowest potential for free-riding and easiest motivated to join. Because of the no free-riding condition of Lemma 3.3, the regulator can include only the \(m\) (or \(m+1\)) largest benefactors into the rescue. If the regulator decides to include bank \(i_{m+1}\) into the bail-in consortium, he/she can avoid the free-riding problem by “burning” an amount equal to \(\frac{1}{\lambda} \left( w_P - w_N - \sum_{j=1}^{m} \nu^{ij} + \nu^{i_{m+1}} \right)\) by giving it away as subsidies. These subsidies can be distributed arbitrarily among banks as long as \(b^i_j - s^i_j \geq \nu^{i_{m+1}}\) for any \(j = 1, \ldots, m+1\). This latter condition ensures that we do not re-introduce the free-riding problem to any other bank: since a bank in the consortium has a welfare-impact of at least \(\nu^{i_{m+1}}\), a rejection of the bank is met by the regulator’s decision to let a default cascade occur.

Remark 3.1. If the regulator had the power to commit to playing \(N\) in the third stage, the equilibrium outcome would improve to \(w_*\) even if \(w_P < w_N\). Commitment power would thus improve social welfare by \((w_P - w_*)1 \{w_P < w_N\}\).

4 Credibility of the regulator’s threat

In this section we identify conditions under which the government threat is credible in the scenario where the regulator is restricted to all-or-nothing rescues. The reason why we perform the credibility analysis for this simpler class of rescues is because the credibility of the threat is the same for all banks, which gives rise to a straightforward characterization. These results complement Theorem 3.5 which characterizes the equilibrium outcome of the game up to the credibility of the regulator’s threat. The first subsection states conditions for a given network, illustrating how the credibility of the regulator’s threat depends on the network topology and the recovery rates. In the sec-
ond subsection, we compare the credibility of the regulator’s threat between two networks. Because the credibility depends on the network topology in a highly nonlinear way, we provide a comparison between two specific network structures that leads to analytically tractable results. We choose the ring network, as a representative structure for sparsely connected networks, and the complete network in representation of densely connected networks.

Because we compare welfare losses in the public bailout to the no-intervention case, we introduce some notation for the sake of brevity. Let $\bar{p}_N$ denote the clearing payment vector in a default cascade and let $\ell_N = \ell(\bar{p}_N)$ and $\delta_N = \delta(\bar{p}_N)$ be the amounts that banks have to liquidate and senior creditors lose, respectively, in a default cascade. Let $D = D(\bar{p}_N)$ denote the set of fundamentally defaulting banks in a default cascade, denote by $C = D \setminus F$ the set of contagious defaults and by $S = D^c$ the set of banks that remain solvent in a default cascade.

### 4.1 Absolute credibility analysis

The main component of welfare losses in a public bailout is the aggregate shortfall $B$ of fundamentally defaulting banks. It can be understood as a measure of the size of the exogenous shock hitting the financial system. The welfare losses in a default cascade $w_N$, on the other hand, is a measure of the shock size after the shock propagates through the financial system. A relevant determinant for the credibility of the regulator’s threat is thus the amplification of the shock through the interbank network. For this analysis, we split the shortfall $B$ into the two components $\|\delta^D_N\|_1$ and $B - \|\delta^D_N\|_1$. Note that only the latter component is amplified through the system as the former component hits the senior creditors of the banks. We define the size of the shock after the amplification as $\|V^C_{0JS} - V^C_{NJS}\|_1 + (1 - \alpha)\|e^F\|_1$, the sum of aggregate losses of non-fundamentally defaulting banks and the liquidation losses of the fundamentally defaulting banks. Our first result states that the regulator’s threat is credible if and only if the amplification of the shock without intervention is smaller than a certain threshold.
Lemma 4.1. Let \( \chi := \| V_0^{C \cup S} - V_N^{C \cup S} \|_1 + (1 - \alpha) \| e^F \|_1 - (B - \| \delta_N^D \|_1) \) denote the amplification of the shock through the interbank network. The regulator’s threat is credible and \( w_N \leq w_P \) if and only if

\[
\chi \leq \lambda (B - \| \delta_N^D \|_1) + \min(\lambda \alpha, 1 - \alpha) \| \ell_R(0) \|_1. \tag{6}
\]

The amplification of the shock depends only on the financial network and it is independent of the regulator’s preference parameter \( \lambda \). Lemma 4.1 thus states that the regulator’s threat becomes more credible if he assigns a larger weight \( \lambda \) to taxpayers’ and senior creditors’ money. Indeed, as \( \lambda \) increases, a bailout is perceived as more costly and the threat to not bail out the fundamentally defaulting banks is more credible.

The amplification of the shock can be decomposed into two components: The first component measures how the initial shock \( B - \| \delta_N^D \|_1 \) spreads through the network and causes losses in interbank assets. Each bank \( i \) is hit by a shock to its interbank assets of size \( \zeta^i := (\pi(L - \bar{p}_N))^i \). The second component consists of the deadweight losses \( (1 - \alpha) \| \ell_N \|_1 \) associated with the liquidation of outside assets after the banks’ interbank assets have been hit by the shock \( \zeta \). The next lemma describes the first component and how it depends on the network topology. Let \( \xi^i = \min(\zeta^i, V_0^i - (1 - \alpha)e^i) \) denote the amount of the shock \( \zeta^i \) that bank \( i \) can absorb.

Lemma 4.2. In any financial system, the aggregate amplification due to network effects equals \( \| \xi^{C \cup S} \|_1 - (B - \| \delta_N^D \|_1) = \frac{1 - \beta}{\beta} \| \bar{p}_N^D \|_1 \). Moreover, if \( \beta < 1 \), then for any set \( S \subseteq C \cup S \) of banks, we have

\[
\zeta^S := \pi^{S,D}(I - \beta \pi^{D,D})^{-1}((1 - \alpha \beta)e^D + (1 - \beta)(e^D + A^D) - V_0^D - \beta \delta_N^D). \tag{7}
\]

The first statement of Lemma 4.2 characterizes the total size of the amplification due to network effects and it gives an easy interpretation of the second term in (1). The second statement helps in obtaining a qualitative understanding for the role of the network topology in the shock’s amplification. If \( \beta < 1 \), one can solve for \( \zeta \) explicitly as in (7) and obtain direct interpre-
tions for the contributing terms. It follows from the definition of $B$ that 
\[ \sum_{i \in D} (1 - \alpha) e^i - V^i_0 = B - \|V^C_0 - (1 - \alpha) e^C\|_1. \]
Equation (7) can thus be understood as follows:

1. The initial shock $B - \|\delta^D_N\|_1$ is increased by the bankruptcy costs $(1 - \beta)\|c^D + A^D + \alpha e^D\|_1$, and the senior creditors’ losses $(1 - \beta)\|\delta^D_N\|_1$, and it is dampened by the available equity that banks in $C$ have after liquidating their outside assets, given by $\|V^C_0 - (1 - \alpha) e^C\|_1$.

2. The shock is non-linearly amplified by the Leontief matrix $(I - \beta \pi^{D,D})^{-1}$ of the subnetwork of defaulting banks $\pi^{D,D}$. A high density of liabilities between banks in $D$ and a low value of $\beta$ make this amplification large.

3. The shock is dispersed among banks in $S$ according to $\pi^{S,D}$. For $S \subseteq S$, a more diversified distribution of liabilities from defaulting to solvent banks reduces deadweight losses from inefficient liquidation of outside assets.

If the interbank asset recovery rate $\beta$ is close to 1, the amplification effects due to the network are small by the first statement of Lemma 4.2. Therefore, the main component of the amplification is due to inefficient liquidation of the outside assets. The following lemma provides a sufficient condition for the regulator’s threat to be credible: if the recovery rate $\alpha$ is sufficiently high ($\alpha > 1/(1 + \lambda)$) or the value of banks’ outside assets make up a sufficiently small proportion of the total value, then the threat is credible for all interbank recovery rates $\beta$ above some threshold.

**Lemma 4.3.** Let $B := \{i \in S \mid (1 + \lambda)(1 - \alpha) e^i > \lambda V^i_0\}$ denote the set of solvent banks, whose value of outside assets accounts for a fraction larger than $\frac{\lambda}{(1 + \lambda)(1 - \alpha)}$ of its total value. Suppose that

\[ (1 - (1 + \lambda)\alpha)\|e^{D\cup B}\|_1 < \lambda \sum_{i \in B \cup C} (e^i + A^i - L^i). \]  

(8)

Then there exists $\beta^* < 1$ such that for all $\beta \geq \beta^*$, the government threat is credible.
The complement of the set $\mathcal{B}$ consists of the banks which satisfy (8) individually. The sufficient condition in Lemma 4.3 can thus be rephrased as follows: All solvent banks which do not satisfy (8) individually have to satisfy (8) in aggregate, when combined with the defaulting banks.

4.2 Relative credibility analysis

In this section, we compare the credibility of the threat between two network topologies and study how this affects the resulting welfare losses. Our first lemma states that equilibrium welfare losses are always lower in networks, in which the regulator’s threat is credible.

Lemma 4.4. For fixed $L, c, e, \alpha, \beta$, the equilibrium welfare losses after intervention are smaller in network $\pi_1$ than in network $\pi_2$ if the regulator’s threat is credible in network $\pi_1$ but not in network $\pi_2$.

Indeed, since the regulator’s threat is credible in network $\pi_1$ but not in network $\pi_2$, Theorem 3.5 implies that $w_{*,1} < w_{N,1} \leq w_P = w_{*,2}$. The next result states that the difference between no-intervention losses and losses in a public bailout are monotonic in the asset recovery rates $\alpha$ and $\beta$. As a consequence, for fixed $\beta$ and $B$, there exists a unique threshold value $\alpha^*$ such that the regulator’s threat is credible for all $\alpha > \alpha^*$. A comparison of credibility between two networks can thus be obtained by comparing the respective values for $\alpha^*$.

Lemma 4.5. For any financial system $(L, \pi, c, e)$, the quantity $w_N - w_P$ is monotonically decreasing in $\alpha$ and $\beta$.

Definition 4.1.

1. We say that the regulator’s threat is more credible in network $\pi_1$ than in network $\pi_2$ for fixed values of $\beta$ and $B$ if $\alpha_{1}^* < \alpha_{2}^*$.

\[^{17}\text{A similar comparison could be obtained by comparing } \beta^*, \text{ the smallest value of the interbank recovery rate such that, for fixed } \alpha \text{ and } B, \text{ the regulator’s threat is credible for any } \beta > \beta^*.\]
2. We say that the regulator’s threat is uniformly more credible in network \( \pi_1 \) than in network \( \pi_2 \) for fixed \( B \) if, for every pair \((\alpha, \beta)\) such that the regulator’s threat is credible in \( \pi_2 \), the threat is also credible in \( \pi_1 \).

To obtain a more quantitative comparison, we choose two specific topologies, which are representative of sparsely and densely connected networks, respectively. We consider the ring and the complete network in our analysis, which is a standard choice in the literature; see also Acemoglu, Ozdaglar and Tahbaz-Salehi (2015). We will assume that aside from the network topology, all parameters are identical so that the difference in the credibility really stems from the pattern of interconnectedness. Specifically, we consider a financial network with \( n \) banks such that after the arrival of the shock, there is precisely one fundamentally defaulting bank, there are \( n_l \) lowly capitalized banks with outside assets \( e_l \) and \( n_h \) highly capitalized banks with outside assets \( e_h > e_l \). Without loss of generality, we assume that bank 1 is the fundamentally defaulting bank, characterized by the value of its outside asset \( e_1 \) and the bailout cost \( B \), which implicitly determine bank 1’s net cash balance \( c_1 \). The value of its outside assets \( e_1 \) may be different from \( e_l \) and \( e_h \). We assume further that all non-fundamentally defaulting banks have an identical net cash balance, that is, \( c_i = c \) for every \( i \neq 1 \) and some constant \( c \), and that all banks have the same total interbank liability, that is, \( L_i = L \) for every bank \( i \) and some constant \( L \). We denote by \( V_{0,l} \) and \( V_{0,h} \) the initial value of a lowly and a highly capitalized bank, respectively.

In the complete network \( \pi_C \), every bank is equally liable to every other bank in the system, i.e., \( \pi_C^{ij} = \frac{1}{n-1} \) for every pair \( i, j \). Interbank liabilities are thus maximally diversified. In a ring network \( \pi_R \), each bank is liable to exactly one other bank so that \( \pi_R^{ij} = 1 \) if and only if \( i = j + 1 \) modulo \( n \). While the complete network is unique, the ring network depends on the labeling of banks. If bank 2 is a highly capitalized bank, it can absorb a larger part of the shock and the resulting welfare losses will be smaller than if bank 2 is a lowly capitalized bank. From an ex-post perspective, the best-possible ring network is the network

\[\text{[18]}\]

The assumption that there is precisely one fundamentally defaulting bank is not crucial for our results, but greatly simplifies the presentation of our results.
where banks $2, \ldots, n_h + 1$ are highly capitalized and the worst-possible ring network is the ring where banks $2, \ldots, n_l + 1$ are lowly capitalized. In this comparison, we will focus on the latter network, where all lowly capitalized banks are hit before the highly capitalized banks. We find that even the worst-possible ring network may outperform the complete network. See panels (b) and (c) in Figure 2 for a graphical representation of the two networks.

Let $B_*$ and $B^*$ denote the thresholds above which all lowly capitalized banks and all highly capitalized banks default in the complete network, respectively. Our next result states that for small shocks $B \leq B^*$ or very large shocks $B > B^*$, the regulator’s threat is uniformly more credible in one network over the other.

**Proposition 4.6.** Fix a shortfall $B$ and suppose that $L \geq \frac{1 + \rho}{\beta} B$, where $\rho = n_l/n_h$. Then the regulator’s threat is uniformly more credible in the complete network if $0 \leq B \leq B_*$ and it is uniformly more credible in the ring network if $B > B^*$.

The condition $L \geq \frac{1 + \rho}{\beta} B$ simply guarantees that interbank liabilities are sufficiently large for the shortfall $B$ to propagate through the system in its entirety. If interbank liabilities were lower, the effects of an increased shortfall $B$ would be felt by the senior creditors but not by the interbank system, thereby not affecting the credibility of the regulator’s threat. Note that the second statement is not vacuous: as illustrated in the examples of the next section, larger shocks may be necessary to cause a systemic default in the ring network than in the complete network.

Proposition 4.6 states that the relative credibility of the regulator’s threat does not depend on the game parameters if the shock size is sufficiently small or sufficiently large. This is different for intermediate shock sizes, where the interbank recovery rate $\beta$ and the size of interbank liabilities $L$ are crucial determinants for the credibility of the threat. We perform a comparative statics analysis for the credibility of the threat along these two dimensions. In both cases, the intuition is the same: if the bankruptcy costs $(1 - \beta)L$ are sufficiently large, the threat is more credible in the ring network.
Proposition 4.7. For $\beta = 1$, there exists $L'$ such that for any $L \geq L'$, the regulator’s threat is more credible in the complete network than in the ring network for any $B \in (B_*, B^*)$. For any $\beta < 1$, there exists $L^*$ such that for any $L \geq L^*$, the regulator’s threat is more credible in the ring network for any $B \in (B_*, B^*)$.

Proposition 4.7 states that unless interbank assets are recovered in full, the regulator’s threat is more credible in the ring network for sufficiently large interbank liabilities. The next proposition shows that if we fix the size of interbank liabilities, the credibility threshold rises from $B_*$ to $B^*$ as the recovery rate increases from a small value to $\beta = 1$.

Proposition 4.8. Suppose that $\beta^{m_l} < \rho$ and $\frac{c^l + \alpha c_l}{c + \alpha c_h} > \frac{1}{n_h}$ and let $L \geq \frac{1+\rho}{\beta} B^*$.

Then there exist $B'(\beta)$ and $B''(\beta)$ with $B_* \leq B' \leq B'' \leq B^*$ such that the following statements hold:

1. The threat is more credible in the complete network for any $B < B'$.
2. The threat is more credible in the ring network for any $B > B''$.
3. $B'$ and $B''$ are monotonically increasing in $\beta$ with $B'(1) = B''(1) = B^*$.

The condition $\beta^{m_l} < \rho$ requires that either the number $n_l$ of lowly capitalized banks in the system is not too small or that the interbank recovery rate is not too high. The condition $\frac{c^l + \alpha c_l}{c + \alpha c_h} > \frac{1}{n_h}$ says that the capital of lowly capitalized banks cannot be too small compared to that of highly capitalized banks. If either condition is violated, the amplification is small in both networks and there may not be a single phase transition. If the conditions are met, there is a nonnegligible amplification of the shock and a phase transition is observed. While it is possible that $B' = B''$, this is not generally true. In the ring network, defaults occur sequentially and the amplification of the shock exhibits discontinuities at shock sizes $B_i$, for which bank $i$ defaults. As a result, the more credible network may switch from the ring network for $B \leq B_i$ to the complete network for $B > B_i$ at these discontinuities. Proposition 4.8 states that any such reversions to the complete network being the more credible
network are local phenomena due to the unequal distribution of the shock’s amplification over the interval \([B_i, B_{i+1}]\). As the shock grows larger, these local effects are dominated by the negative feedback effects of dense interconnections among defaulting banks, making the ring network the more credible network for \(B > B''\). We may interpret the region \(B' < B \leq B''\) as the region where the threat is approximately equally credible in both networks, whereas the complete network is the more credible for \(B < B'\).

Comparing our results to Acemoglu, Ozdaglar and Tahbaz-Salehi (2015), we find that imperfect asset recovery and intervention policies promote sparsely connected networks. While in their model, the complete network is socially preferable for any \(B < B^*\), our analysis shows that for shock sizes \(B \in (B'', B^*)\), more credible bail-in policies are available in the ring network. Moreover, Theorem 3.5 shows that an individual bank’s contribution is larger in the ring network, suggesting that the ring network may be socially preferable where the threat is credible in both networks, even if welfare losses are lower in the complete network without intervention. We illustrate this by means of two examples in the next section.

If we consider an ex-ante perspective where there is uncertainty regarding which bank is going to be hit by a shock, then Proposition 4.8 compares the complete network to the ring architecture for which the shock has the worst impact on the system. Any other ring architecture performs better than the configuration considered. Proposition 4.8 thus provides a lower bound for the credibility of the regulator’s threat in a ring network.

### 5 The impact of network topology on welfare losses

In this section we present numerical examples to illustrate the results stated in our Theorem 3.5 and Propositions 4.6-4.8. In both examples, we compare the ring network to the complete network in a financial system of \(n = 11\) banks with \(\lambda = 1\). An exogenous shock renders bank 1 insolvent and divides the remaining banks into two types: lowly capitalized banks 2,\ldots,6 and highly capitalized banks 7,\ldots,11. We focus mainly on the comparison between the
We compare the densely connected complete network (b) to the sparsely connected ring network (c) for banks with assets and liabilities given in (a). Bank 1 is insolvent, banks 2, ..., 6 are lowly capitalized, and banks 7, ..., 11 are highly capitalized.

The complete network $C$ and the ring network $R$, where all lowly capitalized banks are hit before any of the highly capitalized banks. In Example 1 we will also briefly discuss the ring network $R'$, where highly capitalized banks are hit first. Note that the ring architecture $R$ is the one treated in Propositions 4.6–4.8.

Example 1. Consider a financial network, where banks’ assets and liabilities after the shock are given in the left panel of Figure 2. For the sake of presentation, all interbank liabilities are normalized to 1. In this system, bank 1 is the only fundamentally defaulting bank with a bailout cost of $B = 0.775$. Outside investments can be recovered at $\alpha = 75\%$ of their face value, whereas interbank liabilities are recovered almost in full at $\beta = 90\%$. The three networks considered have relative liability matrices given by $\pi_{ij}^C = \frac{1}{n-1} = 0.1$ for every $i \neq j$, $\pi_{ij}^R = 1$ if and only if $i = j + 1$ (modulo $n$) and $\pi_{ij}^{R'} = 1$ if and only if $i = j - 1$ (modulo $n$).

Observe first that a public bailout is independent of the network topology. In all three networks, a public bailout leads to a welfare loss of $w_P = B + (1 - \alpha)e^1 = 0.85$. The no-intervention outcomes of the three networks are summarized in Table 1, together with the equilibrium welfare losses after...
ter intervention. In the complete network, the shock is spread equally among all banks. This causes the default of lowly capitalized banks, but leaves the highly capitalized banks solvent. In order to remain solvent, however, each highly capitalized bank $i$ has to liquidate an amount $\ell_N^i = 0.24$ of its outside assets. This leads to a welfare loss of 1.01 without intervention, which renders the government threat of no intervention non-credible. As a result, the equilibrium outcome is a public bailout with a welfare loss of 0.85.

In the ring network, the no-intervention outcome depends heavily on the configuration. If the lowly capitalized banks are hit first, the shock tears through their outside assets fairly quickly, leading to their defaults as well as the default of the highly capitalized bank 7. Even though there is one additional default when compared to the complete network, the welfare losses without intervention are smaller in the ring $R$: because of the low capitalization of banks 2, . . . , 6, the resulting liquidation losses are small, and because of the linear network structure, there are no negative feedback effects that cause a decrease in the clearing payments. Even smaller are the welfare losses in the ring configuration $R'$, where only one highly capitalized bank defaults. For both configurations, the regulator’s threat is credible and hence a subsidized bail-in is formed according to Theorem 3.5. The highest-possible incentive-compatible contributions come from the direct creditors of bank 1. In the network $R$, banks 2–5 are each willing to contribute an amount $\nu^i = 0.05$ to a subsidized bail-in, leading to welfare loss of 0.65 after a bail-in. In the network $R'$, banks 10 and 11 contribute $\nu^{10} = 0.19$ and $\nu^{11} = 0.4$, respectively. The resulting welfare loss is 0.26.

**Example 2.** Consider the same financial network as in Example 1 with the

| Network | $\tilde{p}_N$ | $|D|$ | $w_N$ | $w_*$ |
|---------|----------------|-----|-------|-------|
| Complete | (0.05, 0.83, 0.83, 0.83, 0.83, 1.00, 1.00, 1.00, 1.00) | 6 | 1.01 | 0.85 |
| Ring $R$ | (0.13, 0.19, 0.24, 0.29, 0.34, 0.38, 0.94, 1.00, 1.00, 1.00) | 7 | 0.68 | 0.65 |
| Ring $R'$ | (0.13, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 1.00, 0.71) | 2 | 0.48 | 0.26 |

*Table 1:* Shown are the clearing payments, the number of defaulting banks and the welfare losses under no intervention, as well as the equilibrium welfare loss for the three networks.
exception that lowly capitalized banks have a higher amount of outside assets, equal to $e_l = 0.5$. In the complete network, there are no contagious defaults even when there is no intervention. The deadweight losses in the no-intervention case are equal to $w_{N,C} = 0.4667$. In comparison, the ring $R$ leads to the defaults of banks 1, . . . , 3 with a welfare loss of $w_{N,R} = 0.5483$ under no intervention. Nevertheless, the ring network outperforms the complete network under the optimal subsidized bail-in strategy: The maximal incentive-compatible contribution of any bank $i = 2, . . . , 11$ in the complete network is $\xi_i - (1 - \alpha)(\ell^i_N - \ell^i_R(0)) = 0.0583$. The optimal bail-in in the complete network thus involves any seven of the ten banks and results in a welfare loss of $w_{*,C} = 0.4417$. The optimal subsidized bail-in in the ring network involves only banks 2 and 3, but each of them is willing to contribute 0.25, leading to a welfare loss of $w_{*,R} = 0.35$. Because these contributions are much larger than the contribution of any individual bank in the complete network, the resulting welfare losses in a bail-in are smaller than in the complete network, despite the fact that the welfare losses without intervention are smaller in the complete network. This is illustrated graphically in Figure 3.

Observe that in both examples, the systemic default is triggered by a smaller shock in the complete network than in the ring network. In Example 1, all banks default in the complete network for shortfalls larger than $B^* = 3.075$, whereas it takes a shortfall as large as $B_n = 5.7137$ for the last bank in the ring network to default. In the second example, $B^* = 4.575$ and $B_n = 7.7942$. This shows how large the potential for amplification is in a network with dense connections, even at an interbank recovery rate as high as 90%. We re-iterate
the two trade-offs between densely and sparsely connected networks that are illustrated by these examples:

1. Even without any methods of intervention, a higher density of the inter-bank network does not necessarily lead to a reduction of welfare losses. If all banks have a reasonably high level of capitalization as in Example 2, densely connected networks have a potential for absorption of the shock if the shock size is small. If, however, the level of capitalization is low for a large enough fraction of the system as in Example 1, then the dense connections of the defaulting banks cause a negative feedback effect on the available capital for repayments, resulting in an amplification of the shock.

2. In sparsely connected networks, each immediate creditor of the fundamentally defaulting banks is hit by a large part of the shock. The threat of not intervening is thus very severe, creating incentives to contribute large amounts to a bail-in. Figure 3 illustrates that this may lead to lower equilibrium welfare losses, even if the more densely connected network was preferable without intervention.

We conclude this section by showing that these findings are robust to changes in the interbank recovery rate $\beta$ and the size of the ex-post net cash holdings $c^1$ of the fundamentally defaulting bank. All other parameters are taken as in Example 1. A negative value of $c^1$ signifies a net senior debt and $-c^1$ can be interpreted as a measure of the size of the shock. Figure 4a shows the areas in the $(c^1, \beta)$-plane, where the regulator’s threat is more credible in the ring and the complete network, respectively. For the chosen parameters in Example 1, there is a single threshold, where the network with higher credibility switches from the complete to the ring topology.

Figure 4b compares the equilibrium welfare losses in the two networks for a fixed value $\beta = 0.9$ and different values of $\alpha$ and $c^1$. If the recovery rate $\alpha$ is low, the threat fails to be credible in either network and a public

---

20In financial systems with $A^1 = L^1$, such as the systems in Examples 1 and 2, the bailout cost is of the form $B = -c^1 - \alpha e^1$. Thus, $-c^1$ is simply a reparametrization of the bailout cost $B$ that is independent of $\alpha$. 

33
bailout is the only option. For high recovery rates, the threat is credible and a subsidized bail-in can be organized. The steps indicate the contributions of banks to the subsidized bail-in. It is clearly visible that the size of the contributions are much larger in the ring network: we observe 5 small steps, indicating the contributions of lowly capitalized banks to the subsidized bail-in, followed by one large step, the contribution of the first highly capitalized bank in the ring network. For high recovery rates, banks can be incentivized to form a privately-backed bail-in without any contribution of the regulator. In the complete network, banks can be motivated to contribute to a bail-in for smaller shocks than in the ring network, because the threat is more credible for small shocks. The size of the contributions, however, is small.

6 Concluding Remarks

Government support of financial institutions designated to be too big or too important to fail is costly. Various initiatives have been undertaken by central governments and monetary authorities, especially after the global financial cri-

---

Figure 4: The left panel illustrates the regions in the \((c^1, \beta)\)-plane, where the regulator’s threat is more credible in the ring (red) and the complete network (blue), respectively. The border between the two regions is characterized by \(B(c^1) = B_\ast(\beta)\), that is, the more credible network switches at \(B_\ast\) in this case. The threat fails to be credible in either network outside of these regions. The right panel shows equilibrium welfare losses for different recovery rates \(\alpha\) and different net cash holdings \(c^1\). The network with lower welfare losses is socially preferable. The “steps” illustrate the contributions of banks to the bail-in.

\(^{20}\)Figure 4a is obtained as projection of the surfaces \(\alpha_\ast^R\) and \(\alpha_\ast^C\), which intersect precisely in the \(B^\ast\)-plane. Because \(B_\ast\) depends on \(\alpha\), the projection of the intersection is not linear.
sis, to expand resolution plans and tools, including voluntary bail-in consortia contributed by creditors of distressed banks. The central question studied in this paper is: do credible bail-in strategies actually exist? What is the structure of optimal bailouts, and how does it depend on the network structure? We have shown that the existence of credible bail-ins is tightly linked to the amplification of initial shocks through the network. If shocks are strongly amplified by the inefficient liquidation of external assets, bankruptcy costs, and negative feedback effects between interconnected banks in distress, the regulator cannot credibly threaten the banks to not intervene himself. If his no-intervention threat fails to be credible, a public bailout remains the only incentive-compatible rescue option. If the threat is credible, the creditors of the defaulting banks can be incentivized to contribute to a rescue in order to avoid a default cascade.

Our analysis shows that the option of implementing bail-in strategies strongly affects the socially desirable network structure. While in network models without intervention, dense interconnections create a mechanism for the absorption of a shock and, as a result, enhance welfare, these presumptions are reversed if banks and the regulator can strategically coordinate an optimal default resolution plan. Sparser networks enlarge the range of shocks for which a credible bail-in strategy exists; and banks are willing to make larger voluntary contributions to a rescue consortium because they recoup a larger fraction of their social contributions.

Our paper makes a first step towards a systematic analysis of the incentives behind the determination of resolution plans. In a model extension, banks would anticipate which bail-in consortia are credible for which network structures, and hence choose their counterparties so to minimize their ex-ante expected contributions to the equilibrium bail-in plan. While the current analysis indicates that, regardless of the network structure, the threat is credible if recovery rates on external and interbanking assets are high enough, the network structure may play an important role if recovery rates are low. To prevent that banks agree on a network configuration under which the regulator’s threat fails to be credible, the regulator may decide to impose structural
policies which restrict the class of allowable network configurations. Examples of these policies include limiting the size of exposures toward individual counterparties, such as the “Large exposures” framework put forward by the Basel Committee. Such an endogenous network formation model adds an important layer to the moral hazard literature: in addition to maximizing the value of their bailout option through excessive risk taking, as in Farhi and Tirolo (2012), banks can control the likelihood of a public bailout when social losses are high through their interbank lending decisions.

In this study, we have positioned ourselves in an ex-post scenario, i.e., after the realization of asset value shocks. In a future continuation of the work, it would be desirable to account for the (ex-ante) risk taking decisions by banks, which choose their outside investments to maximize shareholder value, while accounting for counterparty risk created by the network. Accounting for ex-ante risk taking behavior will lead to a comprehensive framework for the analysis of welfare maximizing rescue policies. Such a framework would extend existing literature (e.g., Acharya and Yorulmazer (2007) and Acharya, Shin and Yorulmazer (2011)), which has primarily focused on endogenous (non-interbank) asset correlation and liquidity arising when banks make their investment decisions in anticipation of a bailout.

References


21 Under this framework, which is expected to be fully implemented by January 1, 2019, the set of acceptable exposures is computed so as to guarantee that the maximum possible loss incurred by any bank would not be high enough to induce its own default.


A General interventions

A.1 Public bailout

When the regulator is not forced to rescue every bank, he/she will essentially minimize the welfare losses over all possible sets of banks that he/she could bail out. The first lemma describes this minimization procedure.

Lemma A.1. Fix a set of banks \( B \subseteq \{1, \ldots, N\} \). The welfare-maximizing bailout \( s_P(B) \) that ensures solvency of all banks in \( B \) is computed as follows:

1. Let \( p(B) \) be the greatest fixed-point of

\[
p_i = \begin{cases} 
L_i & \text{if } i \in B \text{ or } c^i + \alpha e^i + \sum_{j=1}^n \pi_{ij} p_j \geq L_i, \\
\beta \left( c^i + \alpha e^i + \sum_{j=1}^n \pi_{ij} p_j \right)^+ & \text{otherwise,} 
\end{cases}
\]

and set

\[
s_i(B) := \left( L_i - c^i - \alpha e^i 1_{\{\lambda > 1-\alpha\}} - \sum_{j=1}^n \pi_{ij} p_j(B) \right)^+ 1_{\{i \in B\}}.
\]

2. The optimal bailout is \( s_P(B) = s(\chi(B)) \), where

\[
\chi(B) := \arg\min_{B' \supseteq B} w_\lambda(s(B')).
\]

In words, \( s(B) \) is the welfare-maximizing bailout among all bailouts that rescue \( B \) by giving direct subsidies only to banks in \( B \). The welfare-maximizing bailout \( s_P(B) \) among all bailouts that rescue \( B \) may involve indirect subsidies to rescue banks in \( B^c \) such that banks in \( B \) are protected from secondary effects. The optimal bailout among all bailouts is thus \( s_P(\emptyset) \), where the government is not committed to saving any one bank. Observe that this includes the possibility of not saving any bank when \( \chi(\emptyset) = \emptyset \) and it is thus optimal to let the default cascade play out in full.
Proof of Lemma A.1. Because the greatest fixed point \( p(\mathcal{B}) \) Pareto dominates all other clearing payment vectors in which the set \( \mathcal{B} \) remains solvent, the subsidies \( s(\mathcal{B}) \) are indeed the smallest direct subsidies to banks in \( \mathcal{B} \) required to ensure solvency of banks in \( \mathcal{B} \). To see that it is optimal to give direct subsidies to banks in \( \chi(\mathcal{B}) \), observe that it is clearly optimal to give direct subsidies to the set

\[
\chi'(\mathcal{B}) := \arg \min_{\mathcal{B}' : S(\mathcal{B}') \supseteq \mathcal{B}} w_\lambda(s(\mathcal{B}')), 
\]

where \( S(\mathcal{B}') \) denotes the set of all banks that are solvent after the regulator provides subsidies \( s(\mathcal{B}') \). Clearly, \( S(\chi'(\mathcal{B})) \supseteq \mathcal{B} \) and \( s(\chi'(\mathcal{B})) = s(\chi(\mathcal{B})) \), which establishes the claim. \( \square \)

Corollary A.2. Let \((b, s)\) be the proposed bail-in with response vector \( a \) of the banks. In the last stage, the regulator chooses to implement \( s_P(\emptyset) \) if and only if

\[
w_\lambda(s_P(\emptyset)) \leq w_\lambda(b, s, a) \]

We denote by \( p_* = p(s_P(\emptyset)) \) the clearing payment vector in the optimal bailout (which may include no subsidies at all).

A.2 Equilibrium response of banks

By Corollary A.2 an accepting equilibrium satisfies \( w_\lambda(b, s, a) \leq w_\lambda(s_P(\emptyset)) \). The only accepting equilibrium, where no bank accepts the proposal, arises if the regulator decides to propose the optimal bailout in the first stage.

Lemma A.3. Let \((b, s)\) be a feasible bail-in proposal. In an accepting equilibrium \( a \), bank \( i \) with \( b^i > 0 \) accepts if and only if the following conditions hold:

1. \( w_\lambda(b, s, (0, a^{-i})) \geq w_\lambda(s_P(\emptyset)) \),

2. \( b^i - s^i \leq \sum_{j=1}^n \pi^{ij}(\bar{p}^j(b, s, (1, a^{-i})) - p^j_*) \).

Proof. Fix a feasible proposal \((b, s)\) with accepting equilibrium response \( a \). If Condition 1 is violated, \((1, a^{-i})\) and \((0, a^{-i})\) lead to the same subsidies, the only difference being that \( b^i \) is paid for by bank \( i \) in the former case and by the regulator in the latter. It is thus optimal for bank \( i \) to reject the proposal.
Suppose now that Condition 2 is violated. Feasibility of \((b,s)\) implies that

\[
\ell^i(\bar{p}(b,s,(1,a^{-i}))) = \frac{1}{\alpha} \left( L^i - c^i - s^i + b^i - \sum_j \pi^{ij}(b,s,(1,a^{-i})) \right) +.
\]

Because Condition 2 is violated, it follows that \(\ell^i(\bar{p}(b,s,(1,a^{-i}))) > \ell^i(p_*)\), i.e., bank \(i\) has to liquidate a larger amount of assets in the bail-in than in the bailout. Together with the negation of Condition 2, this implies that \(V^i(\bar{p}(b,s,a)) < V^i(p_*)\). It is thus again optimal to reject the proposal.

Finally, suppose that Conditions 1 and 2 are both met. Condition 1 implies that a rejection will provoke the regulator to resort to the public bailout, and Condition 2 implies that this makes bank \(i\) worse off in the same way as before. It is thus optimal to accept the proposal.

Lemma A.4. Let \((b,s)\) be a proposed bail-in plan with accepting equilibrium responses \(\{a_1, \ldots, a_m\}\). For any \(a_k, k = 1, \ldots, m\), there exists a proposal \((\tilde{b}, \tilde{s})\), to which \(a_k\) is the unique accepting equilibrium response (up to equivalence).

The proof goes analogously to the proof of Lemma 3.4. Note, however, that we do not automatically get Pareto dominance over all rejecting equilibria for a fixed bail-in proposal \((b,s)\). A bank which is not being rescued in \((b,s)\) might be worse off than under \(s_P(\emptyset)\). This does not happen in the optimal bail-in proposal: if the regulator is willing to use taxpayer money to save bank \(i\) in a public bailout, he is willing to pay for it in a subsidized bail-in.

A.3 Optimal proposal of the regulator

To incentivize banks to contribute to a bail-in, the regulator would like to threaten the banks that he/she will let a default cascade occur if they reject the proposal. This threat is not be credible, however, if this leads to higher welfare losses. For a given proposal \((b,s)\), the maximal incentive-compatible contribution of bank \(i\) is \(t^i(b,s) = \left( \sum_j \pi^{ij}(\bar{p}(b,s,1) - p^j_s) \right)^+\) by Lemma A.3. We say that the threat towards bank \(i\) is credible if \(t^i(b,s) > 0\).
Theorem A.5. Fix a set of banks \( B \), set \( \varsigma^i(B) := \sum_{j=1}^n \pi^j_i(p^j(B) - p^j_i) \) and define

\[
\eta^i(B) = \begin{cases} 
\min \left( \varsigma^i(B), \left( c^i + \alpha e^i + (\pi p(B))^i - L^i \right)^+ \right) & \text{if } \lambda \alpha \geq 1 - \alpha, \\
\min \left( \varsigma^i(B), \left( c^i + (\pi p(B))^i - L^i \right)^+ \right) & \text{if } \lambda \alpha < 1 - \alpha.
\end{cases}
\]

Define further \( \nu^i(B) := \lambda \eta^i(B) - (1 - \alpha) \left( \ell_R^i (\eta^i(B) - s^i(B)) - \ell_R^i (-s^i(B)) \right) \). Let \( i_1, \ldots, i_n \) be a decreasing ordering of banks according to \( \nu^i(B) \). Let \( b(B) \) be defined analogous to Theorem 3.5. The welfare-maximizing bail-in \((b_*(B), s_*(B))\) that ensures solvency of all banks in \( B \) is given by \((b(\psi(B)), s(\psi(B)))\), where

\[
\psi(B) := \arg \min_{B' \supseteq B} \min \left( w_\lambda(s(B')) - \sum_{j=1}^{m(B')} \nu^j(B'), w_\lambda(s_F(\emptyset)) - \nu^{m+1}(B') \right),
\]

where \( m(B') = \min \left( k \mid w_\lambda(s(B')) - \sum_{j=1}^k \nu^j(B) < w_\lambda(s_F(\emptyset)) \right) \).

Corollary A.6. The welfare-maximizing bail-in proposal is \((b(\psi(\emptyset)), s(\psi(\emptyset)))\).

The intuition behind Theorem A.5 is exactly the same as in Theorem 3.5, except that the regulator performs an additional minimization procedure over subsets of banks as in Lemma A.1. The regulator can include banks into the bail-in only if the threat towards that bank is credible. Moreover, the regulator will include banks first that have a large benefit from the bail-in proposal. The regulator can include only the first \( m(\psi(B)) \) banks into the bail-in because of the “no free-riding” condition of Lemma A.3. Finally, the welfare-maximizing bail-in proposal is the proposal \((b(\psi(\emptyset)), s(\psi(\emptyset)))\) in Corollary A.6, where the regulator does not commit to saving any bank.

B Proof of Theorems 3.5

We first prove the following preliminary result for the proof of Lemma 3.3

Lemma B.1. Without intervention, the equity value of each bank is \( V^i_N = 0 \) for \( i \in \mathcal{F} \) and \( V^i_N = V^i_0 - \ell^i_N - (1 - \alpha)\ell^i_N \) for \( i \in \mathcal{C} \cup \mathcal{S} \), where \( \ell^i_N = \ell^i(\xi) \).
Proof. Observe that $\xi^i = \zeta^i$ for $i \in S$ and $\xi^i = V_0^i - (1 - \alpha)e^i$ for $i \in C$. We distinguish the two cases. Consider first a bank $i \in S$, liquidating an amount $\ell_N^i = \frac{1}{\alpha}(L^i - z^i - (\pi\bar{p}_N)^i) = \frac{1}{\alpha}(e^i - V_0^i + A^i - (\pi\bar{p}_N))$. The equality $\ell_N^i = \ell^i(\zeta)$ thus follows from $\zeta = \pi(L - \bar{p}_N) = A - \pi\bar{p}_N$. Since $\bar{p}_N^i = L^i$ for $i \in S$, we obtain

$$V^i(\ell_N, \bar{p}_N) = V_0^i - A^i + (\pi\bar{p}_N)^i - (1 - \alpha)e^i = V_0^i - \zeta^i - (1 - \alpha)e^i.$$ 

A bank $i \in C$ liquidates $e^i = \ell^i(V_0 - (1 - \alpha)e)$ and has a value of $V_0^i - (V_0^i - (1 - \alpha)e^i) - (1 - \alpha)e^i = 0$. □

Proof of Lemma 3.3. Fix a feasible proposal $(b, s)$ with accepting equilibrium response $a$. If Condition 1 is violated, $(1, a^{-i})$ and $(0, a^{-i})$ lead to the same subsidies, the only difference being that $b^i$ is paid for by bank $i$ in the former case and by the regulator in the latter. It is thus optimal for bank $i$ to reject the proposal. Suppose now that Condition 2 is violated. Feasibility of $(b, s)$ implies that $\ell_N^i(\bar{p}(b, s, (1, a^{-i}))) = \frac{1}{\alpha}(L^i - c^i - s^i + b^i - \sum_j \pi_{ij}\bar{p}^j(b, s, (1, a^{-i})))$. Because Condition 2 is violated, it follows that $\ell^i(\bar{p}(b, s, (1, a^{-i}))) > \ell^i(s_P)$, i.e., bank $i$ has to liquidate a larger amount of assets in the bail-in than in the bailout. Together with the negation of Condition 2, this implies that $V^i(\bar{p}(b, s, a)) < V_P^i$. It is thus again optimal to reject the proposal.

Finally, suppose that Conditions 1 and 2 are both met. Condition 1 implies that a rejection will provoke the regulator to resort to the public bailout, and Condition 2 implies that this makes bank $i$ worse off in the same way as before. It is thus optimal to accept the proposal. □

Proof of Lemma 3.4. Fix $a_k$ and denote $B := \{i \mid a_k^i = 1\}$ for the sake of brevity. Let $\tilde{b}^i = b^i1_{i \in B}$ and $\tilde{s}^i = s^i$ for $i = 1, \ldots, n$. Because $a_k$ is an accepting equilibrium of $(b, s)$, Conditions 1 and 2 of Lemma 3.3 are satisfied
for any \( i \in B \). It follows again from Lemma 3.3 that \( a_k \) is also an accepting equilibrium of \((\tilde{b}, \tilde{s})\). Because \( b^i = 0 \) for \( i \not\in B \), no bank outside of \( B \) can make any contributions to the bail-in. Condition 1 of Lemma 3.3 implies that if only a subset of \( B \) accepts, the government will not proceed with the bail-in and hence \( B \) is the only accepting equilibrium.

For Pareto-dominance, note that Condition 2 of Lemma 3.3 implies that any \( i \in B \) weakly prefers \( a_k \) over any rejecting equilibrium. Similarly, any \( i \not\in B \) prefers \( a_k \) over any rejecting equilibrium because \( i \) obtains the highest-possible payoff \( V_i^0 \) in \( a_k \). Finally, since \( b^i \geq 0 \) for every \( i = 1, \ldots, n \) and \( w_\lambda(\tilde{b}, \tilde{s}, a_k) = w_P - \lambda \|\tilde{b}\|_1 \), the regulator prefers \( a_k \) over any rejecting equilibrium.

The following lemma formalizes that the regulator cannot gain anything by proposing bail-in plans that will be rejected by any bank. The lemma also rules out the necessity of proposals that award a direct subsidy to banks which have to buy up a positive amount of debt.

**Lemma B.2.** Suppose that \( w_P \geq w_N \). For any proposed bailout \((b, s)\) with equilibrium response \( a \), there exists a proposal \((\tilde{b}, \tilde{s})\) with \( w(\tilde{b}, \tilde{s}, 1) = w(b, s, a) \) and \( \tilde{b}^i \tilde{s}^i = 0 \) for every bank \( i \) such that \( 1 \) is an equilibrium response to \((\tilde{b}, \tilde{s})\).

**Proof.** Let \( A := \{ i \in C \cup S \mid a^i = 0 \} \) denote the set of banks, for which the equilibrium response to \((b, s)\) is to reject the proposal. We define the modified bail-in proposal \((\tilde{b}, \tilde{s})\) as follows: For any bank \( i \in A \), set \( \tilde{b}^i = \tilde{s}^i = 0 \). For any other bank \( i \), set \( \tilde{b}^i := (b^i - s^i)^+ \) and \( \tilde{s}^i := (s^i - b^i)^+ \). Finally, choose \( \tilde{b}^0 \) such that \( \|\tilde{b}\|_1 = \|\tilde{b}\|_1^0 \). The choice of \( \tilde{b}^0 \) ensures that \((\tilde{b}, \tilde{s})\) is feasible. Since \( \tilde{b}^i - \tilde{s}^i = b^i - s^i \) for \( i \in A^c \), it follows that

\[
\tilde{b}^0 + \|\tilde{s}\|_1 = b^0 + \sum_{i=1}^n (b^i - \tilde{b}^i + \tilde{s}^i) = b^0 + \sum_{i \in A} b^i + \sum_{i \in A^c} s^i.
\]

Similarly, \( \ell^c_R(0) = \ell^c(0) \) implies \( \|\ell^{C \cup S}(\tilde{b} - \tilde{s})\|_1 + \|\ell^{F \cup A}_R(0)\|_1 = \|\ell^{A'}(b - s)\|_1 + \|\ell^{F \cup A}(0)\|_1 \) and hence \( w(\tilde{b}, \tilde{s}, 1) = w(b, s, a) \). It remains to check that \( 1 \) is an equilibrium response. Let \( R_{\tilde{b}, \tilde{s}} := \{ i \mid \tilde{b}^i - \tilde{s}^i > 0 \} \) denote the set of banks which have to buy up a positive net amount in \((\tilde{b}, \tilde{s})\). By Lemma 3.3, any
bank $i \not\in \mathcal{R}_{b,s}$ accepts $(\tilde{b}, \tilde{s})$. Moreover, $\tilde{b}^i - \tilde{s}^i = b^i - s^i \leq \xi^i$ for any $i \in \mathcal{R}_{\tilde{b},\tilde{s}}$ because these banks accepted the original proposal $(b, s)$. For any $i \in \mathcal{R}_{\tilde{b},\tilde{s}}$, we compute

$$w(\tilde{b}, \tilde{s}, (0, 1-i)) = w(\tilde{b}, \tilde{s}, 1) + \lambda(\tilde{b}^i - \tilde{s}^i) - (1 - \alpha)(\ell^i(\tilde{b} - \tilde{s}) - \ell^i_R(0))$$

$$= w(b, s, a) + \lambda(b^i - s^i) - (1 - \alpha)(\ell^i(b - s) - \ell^i_R(0))$$

$$= w(b, s, (0, a^{-i})).$$

Since $a$ is an equilibrium response to $(b, s)$, it follows that $w(b, s, (0, a^{-i})) \geq w_N$ from Lemma 3.3. Thus, again by Lemma 3.3, 1 is an equilibrium response to $(b, s)$.

As a consequence of Lemma B.2, we may assume that the optimal bail-in will result from a proposal that is accepted by all banks. By Lemma 3.3, this is the case only if the net contribution of any bank $i$ is smaller than $\xi^i$. The welfare-minimizing bail-in of a fixed consortium $\mathcal{R} \subseteq \mathcal{C} \cup \mathcal{S}$, subject to incentive-compatibility constraints, is thus given by $(b_R, s_R)$ of the following definition.

**Definition B.1.** For a consortium $\mathcal{R} \subseteq \mathcal{C} \cup \mathcal{S}$, let $(b_R, s_R)$ denote the subsidized bail-in defined by

$$b_R^i = \begin{cases} 
\xi^i & \text{if } i \in \mathcal{R}, \alpha \geq \frac{1}{1+\lambda}, \\
\eta^i & \text{if } i \in \mathcal{R}, \alpha < \frac{1}{1+\lambda}, \\
0 & \text{if } i \not\in \mathcal{R},
\end{cases}$$

$$s_R^i = \begin{cases} 
0 & \text{if } \alpha \geq \frac{1}{1+\lambda}, \\
\alpha \ell_R^i(0) & \text{if } \alpha < \frac{1}{1+\lambda},
\end{cases}$$

for $i \in \mathcal{C} \cup \mathcal{S}$ and $b_R^0 = B + \alpha\|e^F\|_1 1\{\alpha<1/(1+\lambda)\} - \|b_{\mathcal{C} \cup \mathcal{S}}^F\|_1$. The net contribution of bank $i \in \mathcal{R}$ to the bail-in $(b_R, s_R)$ increases social welfare by an amount

$$\nu^i := \begin{cases} 
\lambda \xi^i - (1 - \alpha)(\ell_N^i - \ell_R^i(0)) & \text{if } \alpha \geq \frac{1}{1+\lambda}, \\
\lambda \eta^i & \text{if } \alpha < \frac{1}{1+\lambda}.
\end{cases}$$

Let $w_R := w_R(b_R, s_R, 1)$ denote the welfare losses after acceptance by all banks.
Lemma B.3. Fix a rescue consortium $\mathcal{R} \subseteq \mathcal{C} \cup \mathcal{S}$. Among all admissible subsidized bail-ins $(b, s)$ that satisfy $b^i - s^i \leq \xi^i$ as well as $\{i \mid b^i > 0\} \subseteq \mathcal{R}$, the bail-in proposal $(b_\mathcal{R}, s_\mathcal{R})$ minimizes the welfare losses.

Proof. If everybody accepts a proposal $(b, s)$, the subsidized bail-in results in a welfare loss of

$$w(b, s, 1) = \lambda \left( B + \alpha \|e^F\|_1 1_{\{\alpha < 1/(1+\lambda)\}} - \sum_{i \in \mathcal{R}_b} b^i + \sum_{i \notin \mathcal{R}_b} s^i \right) + \frac{1 - \alpha}{\alpha} (b^i + e^i - V_0^i)^+ + \frac{1 - \alpha}{\alpha} (e^i - V_0^i - s^i)^+.$$

If $\alpha \geq 1/(1+\lambda)$, then $(1 - \alpha)/\alpha \leq \lambda$ and hence $w(b, s, 1)$ is non-increasing in $b^i$, $i \in \mathcal{R}_b$ and non-decreasing in $s^i$, $i \notin \mathcal{R}_b$. It follows that the negative welfare impact is minimized at $s = 0$ and the maximum possible $b^i$, for which $(b, s)$ can be accepted, which is $b^i = \xi^i$ for $i \in \mathcal{R}_b$ by Lemma 3.3. If $\alpha < 1/(1+\lambda)$, then $(1 - \alpha)/\alpha > \lambda$. Therefore, $w(b, s, 1)$ is decreasing in $b^i$ on $[0, V_0^i - e^i)$ and increasing in $b^i$ on $(V_0^i - e^i, \xi^i]$. It follows that $w(b, s, 1)$ is minimized at $b^i = \min(\xi^i, (V_0^i - e^i)^+) = \eta^i$ for $i \in \mathcal{R}_b$, where we used that $(V_0^i - e^i)^+ < V_0^i - (1-\alpha)e^i$. Similarly, $w(b, s, 1)$ is decreasing in $s^i$ on $[0, e^i - V_0^i)$ and increasing for $s^i > e^i - V_0^i$. It follows that $w(b, s, 1)$ is minimized at $s^i = (e^i - V_0^i)^+ = \alpha e^i_{\mathcal{R}}(0)$. \hfill \Box

The proof of Theorem 3.5 is concluded by showing that $(b_\mathcal{R}, s_\mathcal{R})$ for $\mathcal{R} = \{i_1, \ldots, i_m\}$ also satisfies condition 1 of Lemma 3.3.

Proof of Theorem 3.5. Consider first the case where $w_P < w_N$ and the regulator’s threat is not credible. If $\alpha \geq 1/(1+\lambda)$, then Lemma 3.3 implies that banks will accept a proposal if and only if they receive a net positive amount, i.e., if $s^i - b^i \geq 0$. In any such proposal, the regulator has to buy up the entire debt $B$, and hence such a proposal is dominated by a public bailout. The regulator will thus resort to a public bailout. If $\alpha < 1/(1+\lambda)$, then Lemma B.3 shows that it is optimal for the regulator to provide a subsidy of $\alpha e^i_{\mathcal{R}}(0)$ to every bank $i \in \mathcal{C} \cup \mathcal{S}$. Since it is necessary that $s^i \geq b^i$ for bank $i$ to accept the
proposal, it follows that the regulator has to cover $B + \alpha \| \ell^F \|_1$ himself. Any such subsidized bail-in leads to a welfare loss of at least $B + \alpha \| \ell_R(0) \|_1 = w_P$. Thus, it is optimal for the regulator to resort to a public bailout.

If $w_N \leq w_P$, then the regulator’s threat is credible. We first show that $w_{C\cup S} \leq w_N$ so that a subsidized bail-in is in the interest of the regulator. The optimal proposal of Lemma B.3 leads to deadweight losses of

$$w_R = \begin{cases} 
\lambda (B - \| \xi^R \|_1) + (1 - \alpha)(\| \ell^R_N \|_1 + \| \ell^R_R(0) \|_1) & \alpha \geq \frac{1}{1+\lambda}, \\
\lambda (B - \| \eta^R \|_1) + \alpha \| \ell_R(0) \|_1 & \alpha < \frac{1}{1+\lambda}.
\end{cases}$$

Thus, for $R = C \cup S$ we obtain $w_{C\cup S} \leq \lambda (B - \| \xi^{C\cup S} \|_1) + (1 - \alpha)(\| \ell_N \|_1)$ with equality if and only if $\alpha \geq 1/(1 + \lambda)$. Together with the first identity in Lemma 4.2 this implies $w_{C\cup S} \leq w_N - (1 + \beta)(1 - \beta)\| A^P - \zeta^P \|_1 \leq w_N$.

We proceed to characterize the optimal proposal of the regulator. Suppose first that $\alpha \geq 1/(1 + \lambda)$. Lemma 3.3 states that a bank $i$ accepts a proposal $(b, s)$ if and only if $b^i - s^i \leq \xi^i$ and $w(b, s, (0, 1^i)) \geq w_N$. Since the optimal choice of $(b, s)$ is $(\xi, 0)$ by Lemma B.3, the latter condition is equivalent to

$$w(b, s, 1) \geq w_N - \nu^i$$

for $\nu^i = \lambda \xi^i - (1 - \alpha)(\ell^i_N - \ell^i_R(0))$. In any proposal, the regulator can thus achieve a welfare loss of $w_N - \min_{i \in R_{b,s}} \nu^i$ at best, where $R_{b,s} = \{ i \mid b^i - s^i > 0 \}$. It is thus in the interest of the regulator to include banks, for which $\nu^i$ is as high as possible. Consider first the proposal $(b_*, s_*)$ with $b_*^i = \nu^i 1\{i \in \{i_1, \ldots, i_k\}\}$ and $s_* = 0$. The definition of $k$ together with Lemma 3.3 imply that this proposal will be accepted by all banks. Since $i_1, i_2, \ldots$ are non-increasingly ordered according to $\nu^i$, it follows that $w_R > w_N$ for any $R \subseteq C \cup S$ with $|R| < k$. Thus, the regulator will not want to propose such a bail-in. Moreover, any $R$ with $|R| = k$ satisfies $w_R \geq w_{\{i_1, \ldots, i_k\}}$, hence the regulator cannot reduce the welfare loss below $w_{\{i_1, \ldots, i_k\}}$ with any such proposal.

If $w_N - w_{\{i_1, \ldots, i_k\}} < \nu^{i_{k+1}}$, then the regulator can improve the proposal $(b_*, s_*)$ to the proposal $(b^*_1, s^*_1)$ with $b^*_1 = \nu^j 1\{i \in \{i_1, \ldots, i_{k+1}\}\}$ and $\| s^*_1 \|_1 = w_N -
By giving away subsidies in the amount of \( w \{ i_1, ..., i_k \} \), the regulator is able to include bank \( i_{k+1} \) into the proposal without violating the condition that \( w(b, s, (0, 1^{-i})) \geq w_N \) for any bank \( i \) of the consortium. Therefore, Lemma 3.3 implies that \((b_\dagger, s_\dagger)\) will be accepted by all banks. Finally, any bail-in \((b, s)\) with \(|\mathcal{R}_{b,s}| > k\) has a lower bound on the welfare loss of \( w_N - \min_{i \in \mathcal{R}_{b,s}} \nu^i \) by Lemma 3.3. Since \( w_N - \min_{i \in \mathcal{R}_b} \nu^i \geq w_N - \nu^{i_{k+1}} \) for any such proposal, it follows that the only reasonable proposals are \((b_*, s_*)\) and \((b_\dagger, s_\dagger)\), of which the regulator will choose whichever minimizes the welfare loss. The proof for \( \alpha < 1/(1 + \lambda) \) works analogously with \( \nu^i = \lambda \eta^i \).

C Proof of Propositions 4.6–4.8

The regulator’s threat is more credible in the ring network than in the complete network if the welfare losses without intervention are smaller in the ring network than in the complete network. The first identity in Lemma 4.2 shows that the welfare losses without intervention are equal to

\[
w_N = (1 - \alpha) \| \ell_N \|_1 + \| \xi_{CUS} \|_1 - B + (1 + \lambda) \| \delta_D \|_1.
\]

Equation (10) gives a convenient way to compute welfare losses once we show that \( L > \frac{1+\rho}{\beta} B \) implies \( \| \delta_D \|_1 = 0 \). We start by characterizing the losses \( \zeta \) to interbank assets in the two networks and the resulting welfare losses. For the sake of brevity, we use \( z_l = V_{0,l} - (1 - \alpha)e_l \) and \( z_h = V_{0,h} - (1 - \alpha)e_h \) to denote the wealth of lowly and highly capitalized banks, respectively, after accounting for the loss due to inefficient liquidation of the assets.

**Lemma C.1.** Fix a shortfall \( B \) of bank 1 and suppose that \( L \geq \frac{1+\rho}{\beta} B \). The shock size \( \zeta_C(B) \) to any bank’s interbank assets in the complete network equals

\[
\zeta_C(B) = \begin{cases} 
\frac{B + (1 - \beta)L}{n-1} & \text{if } B \leq B_*, \\
\frac{B + (n_1 + 1)(1 - \beta)L - n_1 z_l}{n_h + n_1 (1 - \beta)} & \text{otherwise.}
\end{cases}
\]

where \( B_* = (n - 1)z_l - (1 - \beta)L \). All lowly capitalized banks default if and only
if \( B > B_* \) and all banks default if and only if \( B > B^* \), where

\[
B^* = (n_h + n_l(1 - \beta))z_h + n_l(z_l - (1 - \beta)L) - (1 - \beta)L.
\]

For \( B \leq B_* \), the welfare losses without intervention are equal to

\[
w_{N,C}(B) = (1 - \alpha)e^1 + (1 - \beta)L + \frac{1 - \alpha}{\alpha}(B + (1 - \beta)L - (n - 1)c)^+
\]

and for \( B_* < B \leq B^* \), they are equal to

\[
w_{N,C}(B) = (1 - \alpha)e^1 + n_lV_{0,l} + n_h \left( \zeta_C(B) + \frac{1 - \alpha}{\alpha}(\zeta_C(B) - c) \right) - B.
\]

We will be using the following auxiliary result in the proof.

**Lemma C.2.** Let \( \pi \) be an \( m \times m \)-dimensional matrix with diagonal entries of 1 and off-diagonal entries \( a \in \left( \frac{-1}{m-1}, 0 \right) \). The inverse of \( \pi \) has diagonal entries \( b \) and off-diagonal entries \( c \), where

\[
b = \frac{(m - 2)a + 1}{(1 - a)((m - 1)a + 1)}, \quad c = \frac{-a}{(1 - a)((m - 1)a + 1)}. \tag{11}
\]

**Proof.** Let \( B \) denote the \( m \times m \) dimensional matrix with \( b \) on the diagonal and \( c \) elsewhere. Then \( \pi B \) and \( B \pi \) have diagonal entries \( b + (m - 1)ac \) and off-diagonal entries \( c + ab + (m - 2)ac \). It is easy to check that this coincides with the identity matrix if and only if \( b \) and \( c \) are given by (11).

**Proof of Lemma C.1.** Because bank 1 is equally liable to every other bank in the complete network, there are no contagious defaults if and only if the shock \( \zeta_i \) to every bank \( i \neq 1 \) is smaller than the capitalization level \( z_l \) of the lowly capitalized banks. We show that this is precisely the case if \( B \leq B_* \). Indeed, if there are no contagious defaults, every bank \( i \neq 1 \) repays its liabilities in full and bank 1 makes an equilibrium payment in the size of \( \bar{p}_N^{1} = (c^1 + \alpha e^1 + \beta L)^+ = \beta L - B \). Every bank \( i \neq 1 \) is thus hit by a shock to
its interbank assets in the size of
\[
\zeta_i|_{C=\emptyset} = \frac{L - \bar{p}_N^1}{n - 1} = \frac{\min(L, B + (1 - \beta)L)}{n - 1} = \frac{B + (1 - \beta)L}{n - 1}.
\]

(12)

Since we have assumed that there are no contagious defaults, \(\zeta_i|_{C=\emptyset} \leq z_l\) has to hold, which is the case if and only if \(B \leq B_*\). Let \(I_L\) denote the index set of all lowly capitalized banks. Since \(\zeta_i|_{C} \supseteq I_L\ \geq \zeta_i|_{C=\emptyset}\) for any values of \(L\) and \(B\), it follows that \(\zeta_i > z_l\) if and only if \(B > B_*\).

Highly capitalized banks default only if also the lowly capitalized banks default. It is thus necessary that \(B > B_*\). The shock \(\zeta_i\) to interbank assets of a highly capitalized bank is thus given by
\[
\zeta_i|_{C=I_L} = \frac{1}{n-1}(L - \bar{p}_N^1 + n_l(L - \bar{p}_{N,L}))
\]
for
\[
L - \bar{p}_N^1 = \min\left(L, B + (1 - \beta)L + \frac{\beta n_l}{n - 1}(L - \bar{p}_{N,L})\right),
\]
\[
L - \bar{p}_{N,L} = \min\left(L, (1 - \beta)L - z_l + \frac{\beta}{n - 1}(L - \bar{p}_N^1 + n_l(L - \bar{p}_{N,L}))\right).
\]

For the sake of brevity, denote by \(X^1\) and \(X_L\) the values of \(L - \bar{p}_N^1\) and \(L - \bar{p}_{N,L}\), respectively, when these values are strictly smaller than \(L\). We show first that \(\bar{p}_{N,L} \geq \bar{p}_N^1\) with equality if and only if \(\bar{p}_N^1 = \bar{p}_{N,L} = 0\). Suppose that \(\bar{p}_N^1 \geq \bar{p}_{N,L}\), which implies that
\[
X^1 = X_L + B + z_l + \frac{\beta}{n - 1}(\bar{p}_N^1 - \bar{p}_{N,L}) > X_L.
\]

(13)

This leads to a contradiction if \(L > X^1\). Indeed, (13) implies that \(L > X_L\) and hence \(\bar{p}_N^1 = L - X^1 < L - X_L = \bar{p}_{N,L}\), a contradiction. It is thus necessary that \(L - \bar{p}_N^1 = L\) and hence \(\bar{p}_N^1 = 0\) and \(\bar{p}_{N,L} \leq \bar{p}_N^1 = 0\). We have thus shown that either \(L - \bar{p}_{N,L} < L - \bar{p}_N^1\) or \(L - \bar{p}_{N,L} = L - \bar{p}_N^1 = L\).

Since \(X^1\) is bounded above by \(B + \left(1 - \frac{\beta}{1 + \rho}\right)L\), the assumption \(L \geq \frac{1 + \rho}{\beta}B\) implies that \(X^1 \leq L\). In particular, \(L - \bar{p}_N^1 = X^1\) and \(L - \bar{p}_{N,L} = X_L\), which implies that \(\|\delta_N^D\|_1 = 0\). We now use Lemma 4.2 to compute \(\zeta_i|_{C=I_L}\). Since \(\pi^{i,D}\) is an \(n_l + 1\)-dimensional row vector with identical entries \(1/(n - 1)\), it
follows that \((\pi^{i,D}\chi)^j = \sum_k \chi^{k,j}/(n-1)\) for any matrix \(\chi\). Using the notation of Lemma C.2 for \(a = -\beta/(n-1)\) and \(m = |D| = n_t + 1\), each column of \((I - \beta \pi^{D,D})^{-1}\) sums up to
\[
b + n_t c = \frac{1 - a}{1 - a(n_t a + 1)} = \frac{1}{1 + n_t a} = \frac{n - 1}{n - 1 - \beta n_t}.
\]
Since \(n - 1 = n_h + n_t\), each entry of \(\pi^{i,D}(I - \beta \pi^{D,D})^{-1}\) equals \(1/(n_h + n_t(1 - \beta))\) and hence
\[
\zeta^i|_{c = t_L} = \frac{B + (n_t + 1)(1 - \beta)L - n_t z_t}{n_h + n_t(1 - \beta)}.
\]
This shows that the highly capitalized banks default if and only if \(B > B^*\). Welfare losses are now computed easily with (10), recalling that \(\xi^i = \min(\zeta^i, z^i)\) as well as \(V_0^i = (1 - \alpha)e^i + z^i\).

**Lemma C.3.** Fix a shortfall \(B\) of bank 1 and suppose that \(L \geq \frac{1}{3}B\). Bank \(i+1\) defaults in the ring network if and only if \(B > B_{i+1}\), where \(B_2 = z_t - (1 - \beta)L\) and
\[
B_{i+1} = \max \left( B_2, \frac{2 - \beta - \beta^{i-1}}{\beta^{i-1}(1 - \beta)} z_t + \frac{2 - \beta - \beta^{i-n_t-1}}{\beta^{i-1}(1 - \beta)} (z_h - z_t) 1_{\{i > n_t\}} - \frac{1 - \beta^i}{\beta^{i-1}L} \right)
\]
for \(i > 2\). The welfare losses without intervention are equal to
\[
w_{N,R}(B_{i+1}) = (1 - \alpha)e_1 + iV_{0,t} + (i - n_t)^+ (e_h - e_t) - B_{i+1}.
\]

**Proof.** Since \(L \geq B + (1 - \beta)L\) by assumption, the shock size \(\zeta^2\) to bank 2’s interbank assets is equal to \(\zeta^2 = B + (1 - \beta)L\), which implies \(B_2 = z_t - (1 - \beta)L\). We will derive \(B_{i+1}\) for \(i > 1\) recursively. Any bank \(i + 1\) defaults if and only if all banks \(2, \ldots, i\) default and the shock to bank \(i + 1\)’s interbank assets exceeds \(z^{i+1}\), that is, \(\zeta^{k+1} > z^{k+1}\) for \(k = 1, \ldots, i\). In particular, this shows that \(B_{i+1} \geq B_2\) for any \(i > 1\). The recursive relation
\[
\zeta^{k+1} = \min(L, (1 - \beta)L + \beta \zeta^k - z^k)^+)
\]
implies that bank \(k\)’s default leads to a minimal shock to bank \(k+1\)’s interbank
assets of size \((1 - \beta)(L - z^k)\). If this minimal shock size is larger than \(z^{k+1}\), bank \(k + 1\) defaults at the same time as bank \(k\) and \(\zeta^{k+1} \geq \zeta^k\). If \((1 - \beta)(L - z^k) < z_{k+1}\), then \(\zeta^{k+1} < \zeta^k\) and hence \(B_{k+1} \geq B_k\).

Consider first the case where \(i \leq n_l\). If \(\beta \leq 1 - \frac{z_l}{L - z_l}\), then \((1 - \beta)(L - z_l) \geq z_l\) and hence the default of any lowly capitalized bank \(k > 2\) occurs at the same time as the default of bank \(k\). It follows that \(B_{i+1} = B_2\) in that case.

If \(\beta > 1 - \frac{z_l}{L - z_l}\), then \((\zeta^k)_{k \geq 2}\) is decreasing in \(k\). In particular, \(L \geq \frac{1}{\beta}B\) implies that \(L \geq \zeta^2 \geq \zeta^k\) and hence we can find \(\zeta^{i+1}\) with the recursion \(\zeta^{k+1} = ((1 - \beta)L + \beta \zeta^k - z_l)^+\), starting from \(\zeta^2 = \min(L, B + (1 - \beta)L)\). A short calculation shows that

\[
\zeta^{i+1} = \left(\beta^{i-1}B + (1 - \beta^i)L - \frac{1 - \beta^{i-1}}{1 - \beta}z_l \right)^+ 
\]  

for \(i \leq n_l + 1\). Bank \(i + 1 \leq n_l + 1\) defaults if and only if \(\zeta^{i+1} > z_l\). Solving \(\zeta^{i+1} = z_l\) for \(B\) yields \(B_{i+1}\) in the case \(i \leq n_l\).

Consider now the case where \(i > n_l\). Similarly as before, if \(\beta \leq 1 - \frac{z_h}{L - z_h}\), then \(B_{i+1} = B_2\) for every bank \(i + 1\). Suppose, therefore, that \(\beta > 1 - \frac{z_h}{L - z_h}\) and hence \((\zeta^k)_{k \geq n_l + 2}\) is decreasing in \(k\). For \(k \leq n_l + 2\), equation (16) implies that \(\zeta^k < L\) if and only if \(L > \frac{1}{\beta} - \frac{1 - \beta^{k-2}}{1 - \beta}z_l\). In particular, \(\zeta^k \leq L\) for any \(k\).

It is thus sufficient to consider the recursion \(\zeta^{k+1} = ((1 - \beta)L + \beta \zeta^k - z_l)^+\) with the explicit solution

\[
\zeta^{i+1} = \left(\beta^{i-1}B + (1 - \beta^i)L - \frac{\beta^{i-n_l-1} - \beta^{i-1}}{1 - \beta}z_l - \frac{1 - \beta^{i-n_l-1}}{1 - \beta}z_h \right)^+. 
\]  

Solving for \(\zeta^{i+1} = z_h\) yields \(B_{i+1}\) in the case \(i > n_l\). The welfare losses without intervention follow directly from \((10)\).

We next present two auxiliary lemmas that will be needed in the proof of Proposition 4.8.

**Lemma C.4.** Suppose that \(i > j \geq 0\) and \(i \geq 2\). Then

\[
g_{i,j}(\beta) := \frac{2 - \beta}{1 - \beta} - \frac{\beta + j(1 - \beta)\beta^{i-j-1} - \beta^{i-j}}{(i - 1)(1 - \beta)^2}.
\]  

\[52\]
is increasing in $\beta$ with $\lim_{\beta \to 1} g_{i,j}(\beta) = 1 + \frac{(i-j-1)(i+j)}{2(i-1)}$.

Proof. We show monotonicity by induction on $i = j + k$. The result is trivial for $i = j + 1$ since $g_{i,j} \equiv 1$ in that case. Observe that $g_{j+k,j}$ satisfies the recursive identity

$$g_{j+k+1,j}(\beta) = 2 + \frac{j+k-1}{j+k} \beta (g_{j+k,j}(\beta) - 1).$$

Taking the derivative with respect to $\beta$ we obtain

$$\frac{\partial g_{j+k+1,j}(\beta)}{\partial \beta} = \frac{j+k-1}{j+k} \left( g_{j+k,j}(\beta) - 1 + \beta \frac{\partial g_{j+k,j}(\beta)}{\partial \beta} \right).$$

Inductively, we derive that $g_{j+k+1,j}(\beta) \geq 1$ for any $\beta$ and hence $\frac{\partial g_{j+k+1,j}(\beta)}{\partial \beta} \geq 0$.

The limit result is a straight-forward application of L'Hôpital's rule,

$$\lim_{\beta \to 1} g_{i,j}(\beta) = \lim_{\beta \to 1} \frac{(i-1)(2\beta-3) - 1 + (j+1)(i-j)\beta^{i-j-1} - j(i-j-1)\beta^{i-j-2}}{-2(i-1)(1-\beta)}$$

$$= 1 + \frac{(i-j-1)(i+j)}{2(i-1)}.$$

**Lemma C.5.** If $L \geq i z_l + (i - n_l)^+(z_h - z_l)$, then $B_{i+1} + (1 - \beta)L$ is non-decreasing in $\beta$, where $B_{i+1}$ is given in Lemma C.3.

Proof. The statement holds trivially where $B_{i+1} + (1 - \beta)L = z_l$. To see that the quantity $B_{i+1} + (1 - \beta)L$ is non-decreasing everywhere, suppose first that $i \leq n_l$ and take the derivative to obtain

$$\frac{\partial (B_{i+1} + (1 - \beta)L)}{\partial \beta} = \frac{i-1}{\beta(i)} (L - g_{i,0}(\beta)z_l).$$

Since $g_{i,0}$ is increasing by Lemma C.4 with $g_{i,0}(1) = 1 + \frac{i}{2}$, the assumption $L \geq i z_l$ implies that $L \geq g_{i,0}(1)z_l \geq g_{i,0}(\beta)z_l$ and hence the derivative is positive everywhere. For $i > n_l$, a straight-forward computation shows that

$$\frac{\partial B_{i+1} + (1 - \beta)L}{\partial \beta} = \frac{i-1}{\beta(i)} (L - g_{i,0}(\beta)z_l - g_{i,n_l}(\beta)(z_h - z_l)).$$

53
The condition on $L$ now implies that the derivative is positive everywhere in the same way as before.

**Proof of Proposition 4.6.** We consider first the case where $B \leq B_*$. We will show that $w_{N,R}(B) \geq w_{N,C}(B)$ for any values of $\alpha$ and $\beta$, which shows that the threat is uniformly more credible in the complete network than in the ring network. If bank 1’s shortfall $B$ is lower than $(n-1)c - (1-\beta)L$, then no banks other than bank 1 have to liquidate anything in the complete network. Thus, $w_{N,C}(B) = (1-\alpha)e^1 + (1-\beta)L$ attains the minimum possible value for welfare losses and hence $w_{N,C}(B) \leq w_{N,R}(B)$. Suppose next that $B > (n-1)c - (1-\beta)L$ and that $B = B_{i+1}$ for some bank $i$. Lemmas C.1 and C.3 imply that $w_{N,R}(B_{i+1}) \geq w_{N,C}(B_{i+1})$ if and only if

$$iz_l + (i-n_l)^+(z_h - z_l) + (n-i-1)(1-\alpha)c \geq B_{i+1} + (1-\beta)L.$$  

(18)

This inequality is clearly satisfied if $B_{i+1} + (1-\beta)L \leq iz_l + (i-n_l)^+(z_h - z_l)$. Suppose, therefore, that $B_{i+1} + (1-\beta)L > iz_l + (i-n_l)^+(z_h - z_l)$ holds. The condition $L > \frac{1-\beta}{\beta}B_{i+1}$ implies $L > B_{i+1} + (1-\beta)L > iz_l + (i-n_l)^+(z_h - z_l)$, hence Lemma C.5 shows that the right-hand side of (18) is increasing in $\beta$. The maximum is thus attained at $\beta = 1$. Lemma C.3 shows that $B_{i+1}$ converges to $iz_l + (i-n_l)^+(z_h - z_l)$ as $\beta \to 1$, which satisfies (18). It follows that $w_{N,R}(B_{i+1}) \geq w_{N,C}(B_{i+1})$ for any values of $\alpha, \beta$ and $i$. To see that the statement also holds for $B \in (B_i, B_{i+1}]$, note that

$$w_{N,R}(B) = w_{N,R}(B_i) + (1-\alpha)e_N^{i+1} + \zeta^{i+1} - (B - B_i).$$

It follows from (17) that

$$\frac{\partial w_{N,R}(B)}{\partial B} = \begin{cases} \beta^{i-1} - 1 & \text{if } B_{i+1} < B \leq B_i + (c - (1-\beta)(L-z_l))^+; \\ \frac{\beta^{i-1} - \alpha}{\alpha} & \text{otherwise.} \end{cases}$$  

(19)

Lemma C.1 implies that $\frac{\partial w_{N,C}(B)}{\partial B} = \frac{1-\alpha}{\alpha}$, showing that the rate of increase of welfare losses is larger in the complete network than in the ring network.
Since we have shown $w_{N,C}(B_{i+1}) \leq w_{N,R}(B_{i+1})$ already, this shows that also $w_{N,C}(B) \leq w_{N,R}(B)$ for any $B \in (B_i, B_{i+1}]$. Finally, for any $B > B^*$, all banks default in the complete network, implying that $w_{N,C}(B) \geq w_{N,R}(B)$. □

Lemma C.6. If $L \geq \frac{1}{\beta} B_s$, then $B_{n_l+1} < B_s$.

Proof. Suppose towards a contradiction that $B_{n_l+1} \geq B_s$. Then $L \geq \frac{1}{\beta} B_s$ implies that $L \geq n_l z_l$, hence Lemma C.5 shows that the maximum of $B_{n_l+1} + (1 - \beta)L$ is attained at $\beta = 1$. Since $B_{n_l+1} \rightarrow n_l z_l$ as $\beta \rightarrow 1$, this yields

\[ B_{n_l+1} + (1 - \beta)L \leq n_l z_l < (n - 1)z_l = B_s + (1 - \beta)L, \]

which is a contradiction. □

Proof of Proposition 4.7. Consider first the case where $\beta = 1$. A straightforward application of Lemmas C.1 and C.3 yields that

\[ w_{N,R}(B_{i+1}) - w_{N,C}(B_{i+1}) = (i - n_l)V_{0,h} + \frac{1 - \alpha}{\alpha} n_h c - \frac{n_h}{\alpha} \zeta_C(B_{i+1}) \]

\[ = \frac{1 - \alpha}{\alpha} (n - 1 - i)c \]

for any $i$ with $B_{i+1} \geq B_s$. This implies that $w_{N,R}(B_{i+1}) \geq w_{N,C}(B_{i+1})$ in the same way as in the proof of Proposition 4.6. For the second statement where $\beta < 1$, we will show that $w_{N,C}(B) \geq w_{N,R}(B)$ holds for $L$ above the threshold

\[ L^* := z_h + \frac{z_h - z_l}{\rho(1 - \beta)}. \]  

Let $B \in (B_i, B_{i+1}]$ for some $i$. Since $B > B_s$, Lemma C.6 implies that $i > n_l$. Monotonicity of $w_{N,R}(B)$ and $w_{N,C}(B)$ in $B$ together with Lemmas C.1 and C.3
show that

\[ w_{N,R}(B) - w_{N,C}(B) \leq w_{N,R}(B) - \lim_{B \searrow B^*} w_{N,C}(B) \]
\[ = (i - n_I) V_{0,h} + \frac{n_h}{\alpha d} (1 - \alpha) dc - n_h z_l - n_I (1 - \beta) L \]
\[ \leq \frac{n_h}{\alpha d} (dz_h - n_h z_l - n_I (1 - \beta) L) \]
\[ \leq \frac{n_h}{\alpha d} (n_h (z_h - z_l) - n_I (1 - \beta) (L - z_h)). \]

This bound is smaller or equal to 0 if \( L \geq L^* \), which shows the claim.

\[ \square \]

**Proof of Proposition 4.8.** Observe first that the result holds true by Propositions 4.6 and 4.7 if \( \beta = 1 \), stating that \( B' = B'' = B^* \), or if \( \beta < 1 \) and \( L \geq L^* \), stating that \( B' = B'' = B^* \), where \( L^* \) is given in (20). Suppose, therefore, that \( \beta < 1 \) and \( L < L^* \). Let \( i_0 \) denote the smallest integer \( i \) such that \( B_{i+1} > B^* \) and note that \( i_0 > n_I \) by Lemma C.6. Consider the sequence \( Z := (Z_{i+1})_{i \geq i_0} \) of differences \( Z_{i+1} := w_{N,R}(B_{i+1}) - w_{N,C}(B_{i+1}) \) between welfare losses in the two networks. We will show that \( Z = (Z_{i+1})_{i \geq i_0} \) is non-increasing under the stated assumptions. Lemmas C.1 and C.3 imply that

\[ w_{N,R}(B_{i+1}) - w_{N,R}(B_i) = V_{0,h} - (B_{i+1} - B_i), \]
\[ w_{N,C}(B_{i+1}) - w_{N,C}(B_i) = \left( \frac{n_h}{\alpha d} - 1 \right) (B_{i+1} - B_i), \]

where we denote \( d = n_h + n_I (1 - \beta) \) for the sake of brevity. Let \( \hat{\beta}_i \) be the solution to \( V_{0,h} = \frac{n_h}{\alpha d} (B_{i+1} - B_i) \), that is, where \( Z_{i+1} - Z_i = 0 \). We first show that \( (\hat{\beta}_i)_{i \geq i_0} \) is decreasing. Indeed, observe that

\[ B_{i+1} - B_i = \frac{1}{\beta^{i-1}} (z_h - (1 - \beta) (L - z_h)). \]

Since \( B_{i+1} \geq B^* \), it is necessary that \( z_h > (1 - \beta) (L - z_h) \) as otherwise \( B_{i+1} \) would equal \( B_2 < B^* \) by Lemma C.3. It follows that \( B_{i+1} - B_i \) is increasing.
in $i$. Taking the derivative of $B_{i+1} - B_i$ with respect to $\beta$ yields

$$\frac{\partial (B_{i+1} - B_i)}{\partial \beta} = \beta + (i - 1)(1 - \beta) \left( L - z_h - \frac{i - 1}{\beta + (i - 1)(1 - \beta)} z_h \right).$$

Since $L > B^* > (n - 1) z_h$, this shows that $B_{i+1} - B_i$ is increasing in $\beta$ for any $i > n_l$. Using these monotonicity properties and the definition of $\hat{\beta}_i$, we obtain

$$V_{0,h} = \frac{n_h}{\alpha d} \left( B_{i+1}(\hat{\beta}_i) - B_i(\hat{\beta}_i) \right) < \frac{n_h}{\alpha d} \left( B_{i+2}(\hat{\beta}_i) - B_{i+1}(\hat{\beta}_i) \right).$$

Since $B_{i+2} - B_{i+1}$ is increasing in $\beta$, this shows that $\hat{\beta}_{i+1} < \hat{\beta}_i$. For a fixed $\beta$, it follows that $Z_{i+1} - Z_i$ is positive if and only if $\beta < \hat{\beta}_i$. Since $(\hat{\beta}_i)_{i \geq i_0}$ is decreasing in $i$, it follows that $Z_{i+1} - Z_i$ is positive for all $i \leq i^*$, where $i^* = \max \{k \mid \beta < \hat{\beta}_k\}$, and non-positive for all $i > i^*$. It remains to show that $i^* < i_0$. The indices $i^*$ and $i_0$ are the unique integers $i$ and $j$, respectively, for which $\hat{\beta}_{i+1} \leq \beta < \hat{\beta}_i$ and $B_j \leq B_i < B_{j+1}$. Thus, the definition of $\hat{\beta}_i$ and the expression for $B_j$ in Lemma C.3 imply that $i^*$ and $i_0$ are defined by

$$\frac{1}{\beta_{i^* - 1}} < \frac{\alpha d V_{0,h}}{n_h(z_h - (1 - \beta)(L - z_h))} \leq \frac{1}{\beta_{i^*}}, \quad \frac{1}{\beta_{i_0 - 2}} \leq h(\beta) < \frac{1}{\beta_{i_0 - 1}},$$

(21)

where

$$h(\beta) = \frac{z_h - (1 - \beta) z_l + (1 - \beta) \beta^n ((n - 1) z_l - L)}{\beta^n (z_h - (1 - \beta) L - z_h)}. $$

Because $\alpha V_{0,h} = \alpha + \alpha e_h \leq c + \alpha e_h = z_h$, it is sufficient to show that

$$h(\beta) \geq \frac{dz_h}{n_h(z_h - (1 - \beta)(L - z_h))}$$

(22)

as this readily implies $\frac{1}{\beta_{i^* - 1}} < \frac{1}{\beta_{i_0 - 1}}$ by (21), which is equivalent to $i^* < i_0$. A short computation shows that (22) is equivalent to

$$L \leq \frac{z_h - z_l}{\beta^n (1 - \beta)} + (n - 1) z_l - \rho z_h + \frac{n_h z_l - z_h}{n_h(1 - \beta)}. $$

(23)

57
The conditions \( \beta^m < \rho \) and \( \frac{z_i}{z_h} > \frac{\rho}{n-1} \) together with \( L < L^* \) imply that

\[
L < \frac{z_h - z_l}{\rho(1 - \beta)} + z_h \leq \frac{z_h - z_l}{\beta^m(1 - \beta)} + (n - 1)z_l - \rho z_h.
\]

This implies (23) because the last term in (23) is non-negative by assumption.

Let \( j \) denote the largest integer \( k \) such that \( w_{N,R}(B_{k+1}) \geq w_{N,C}(B_{k+1}) \). Fix a shortfall \( B \in (B_i, B_{i+1}] \) for some \( i \) with \( i_0 \leq i \leq j \). It follows from (19) and

\[
\frac{\partial w_{N,C}(B)}{\partial B} = \frac{n_h}{\alpha d} - 1
\]

that \( \frac{\partial w_{N,C}(B)}{\partial B} \geq \frac{\partial w_{N,R}(B)}{\partial B} \) if and only if \( f_i(\beta) \geq \frac{m_i}{n_h} \), where \( f_i(\beta) = \frac{1 - \beta^{i-1}}{\beta^i(1 - \beta)} \).

A short computation shows that \( f_i \) is decreasing in \( \beta \) with \( \lim_{\beta \to 1} f_i(\beta) = i - 1 \).

It follows that \( f_i(\beta) > f_i(1) = i - 1 \geq n_l \), which implies \( \frac{\partial w_{N,C}(B)}{\partial B} > \frac{\partial w_{N,R}(B)}{\partial B} > 0 \).

Since \( w_{N,R}(B_{i+1}) \geq w_{N,C}(B_{i+1}) \) by the choice of \( i \) and the infinitesimal increase of \( w_{N,C} \) is larger than that of \( w_{N,R} \) on the entire interval \((B_i, B_{i+1}]\), this shows that \( w_{N,R}(B) \geq w_{N,C}(B) \) for any \( B \in (B_i, B_{i+1}] \). In particular, we have shown that \( w_{N,R}(B) \geq w_{N,C}(B) \) for any \( B \in (B_s, B_{j+1}] \) and hence \( B_{j+1} \) serves as \( B' \).

Similarly, we observe that \( w_{N,C}(B) - w_{N,R}(B) \) remains positive for \( B > B_k \) where \( B_k \) is the smallest integer \( i \), for which \( \lim_{B \searrow B_k} w_{N,C}(\tilde{B}) - w_{N,R}(\tilde{B}) = 0 \).

Indeed, on \((B_s, B^*)\), \( w_{N,C} \) is continuous and \( w_{N,R} \) has discontinuities only at \( B_i \) with a jump size \( \Delta w_{N,R}(B_{i+1}) = (1 - \beta)(L - z_h) \) for any \( i > n_l + 1 \).

Since the jump sizes are constant and \((Z_{i+1})_{i \geq i_0} \) is decreasing, it follows that \( \lim_{B \searrow B_i} w_{N,C}(\tilde{B}) - w_{N,R}(\tilde{B}) = 0 \) for all \( i \geq k \).

Since \( \frac{\partial w_{N,C}(B)}{\partial B} - \frac{\partial w_{N,R}(B)}{\partial B} > 0 \) on \((B_i, B_{i+1}]\) for any \( i \geq k \) and \( w_{N,C} - w_{N,R} \) is positive at left limits of these intervals, it follows that \( w_{N,C} \geq w_{N,R} \) on \([B_k, B^*]\) and hence \( B_k \) serves as \( B'' \).

Finally, let \( j_0 \) denote the largest integer such that \( B_{j_0+1} \leq B^* \). It follows from the definition of \( B^* \) and Lemma C.1 that \( w_{N,C}(B^*) + B^* \) is constant in \( \beta \).

Equation (19) shows that \( \frac{\partial w_{N,R}(B) + B}{\partial B} \) is positive on intervals \((B_i, B_{i+1}]\). Since \( \frac{\partial B^*}{\partial \beta} \) is positive as well, it follows that \( w_{N,R}(B^*) - w_{N,C}(B^*) = w_{N,R}(B^*) + B^* - (w_{N,C}(B^*) + B^*) \) is increasing in \( \beta \), where \( B_{j_0+1} \) is constant. Because \( Z_{i+1} - Z_i = V_{0,h} - \frac{n_h}{\alpha d} (B_{i+1} - B_i) \) is decreasing in \( \beta \), we deduce by backward induction that \( Z_{i+1} \) is increasing in \( \beta \) for any \( i \) with \( i_0 \leq i \leq j_0 \). This implies...
that the sequence \((Z_i)_{i \geq i_0}\) crosses the thresholds 0 and \(-(1 - \beta)(L - z_h)\) later as \(\beta\) increases and hence \(B_{j+1}\) and \(B_k\) are increasing in \(\beta\).

\[\square\]

D Proofs of auxiliary results

Lemma 2.1 follows as a consequence of the following result, which is a straightforward adaptation of Theorem 1 in Rogers and Veraart (2013) to our setting.

**Lemma D.1.** Let \(\varphi^{(k)}\) denote the \(k\)-fold application of the operator

\[
\varphi^i(p) := \begin{cases} 
L^i & \text{if } c^i + \alpha e^i + \sum_{j=1}^{n} \pi_{ij} p^j \geq L^i, \\
\beta \left( c^i + \alpha e^i + \sum_{j=1}^{n} \pi_{ij} p^j \right)^+ & \text{otherwise.}
\end{cases}
\]

Then \(\bar{p} = \lim_{k \to \infty} \varphi^{(k)}(L)\).

**Simultaneous proofs of Lemma 2.1 and D.1.** Any clearing payment vector \(p\) is a fixed point of the operator \(\varphi\) because a bank \(i\) defaults if and only if it cannot repay its liabilities after the liquidation of the entire outside asset. The same arguments as in the proof of Theorem 1 in Rogers and Veraart (2013) show that there exist \(\bar{p}\) and \(\bar{p}\) such that \(p \leq p \leq \bar{p}\) for any fixed point \(p\) of \(\varphi\). Moreover, \(\bar{p}\) can be found with the algorithm in Lemma D.1. The final statement follows from monotonicity of the banks’ value of equity and the deadweight losses in the clearing payment vector \(p\). \(\square\)

**Proof of Lemma 4.1.** By the definition of a clearing payment vector, we have

\[\bar{p}_N - \beta \delta_N = \beta c^i + \alpha \bar{e}^i + \beta (\pi \bar{p}_N)^i \text{ for any } i \in D,\]

which can be rewritten as

\[V_0^i - (1 - \alpha \beta) e^i - (1 - \beta)(c^i + A^i) - \beta \zeta^i + L^i - \bar{p}_N^i + \beta \delta_N^i = 0, \quad i \in D. \tag{24}\]

Since \(\bar{p}_N^i = L^i\), summing \(L^i - \bar{p}_N^i\) over \(i \in D\) yields \(\|L - \bar{p}_N\|_1\). Because \(\pi\) is doubly stochastic, this is equal to \(\|\zeta\|_1\). Summing (24) over \(i \in D\) thus yields

\[\|V_0^C - (1 - \alpha) e^C\|_1 + \|\zeta^S\|_1 - B + \beta \|\delta_N^D\|_1 - (1 - \beta) \sum_{i \in D} (c^i + \alpha e^i + A^i - \zeta^i) = 0.\]

Using the definition of \(\xi\), we can rewrite this identity as

59
\[ \| \xi^{C \cup S} \|_1 - (B - \| \delta^P_N \|_1) = \frac{1 - \beta}{\beta} \| \bar{p}_N \|_1, \] (25)

hence \( w_N = (1 - \alpha) \| \ell_N \|_1 + \| \xi^{C \cup S} \|_1 - B + (1 + \lambda) \| \delta^P_N \|_1. \) For \( i \in C \cup S, \) we have \( \xi^i + (1 - \alpha) \ell^i_N = V^i_0 - V^i_N \) by Lemma B.1. Now \( B = \sum_{i \in F} (1 - \alpha) e^i - V^i_0 \) implies
\[ w_N = \| V^0_{C \cup S} - V^F_{C \cup S} \|_1 + (1 - \alpha) \| e^F \|_1 - B + (1 + \lambda) \| \delta^P_N \|_1. \]
Together with Lemma 3.1, this shows that \( w_N - w_P \leq 0 \) is equivalent to (6).

**Proof of Lemma 4.2.** We have shown the first statement in (25) already. For the second statement, observe that the clearing payment \( \bar{p}^D_N \) is a solution to
\[ \frac{1}{\beta} p^D_N - \delta^P_N = c^D + \alpha e^D + \pi^D S L^S + \pi^D R \bar{p}^D_N. \] (26)

For \( \beta < 1 \), the spectral radius of \( I - \beta \pi^D R \) is strictly smaller than 1 and hence \( I - \beta \pi^D R \) is invertible. The system (26) thus admits the explicit solution
\[ \bar{p}^D_N = \beta \left( I - \beta \pi^D R \right)^{-1} \left( c^D + \alpha e^D + \pi^D S L^S + \delta^P_N \right). \]
Using that \( L^D = \left( I - \beta \pi^D R \right)^{-1} \left( L^D - \beta \pi^D R L^D \right) \) and \( A^D = \pi^D R L^D + \pi^D S L^S \), we obtain
\[ L^D - \bar{p}^D_N = \left( I - \beta \pi^D R \right)^{-1} \left( L^D - \beta e^D - \alpha \beta e^D - \beta A^D - \beta \delta^P_N \right). \]
This readily implies the statement. \( \square \)

**Proof of Lemma 4.3.** Equation (10) implies that \( w_N \leq w_P \) if and only if
\[ (1 + \lambda) \frac{1 - \beta}{\beta} \| \bar{p}_N \|_1 \leq \lambda \| \xi^{C \cup S} \|_1 - (1 - \alpha) \| \ell_N \|_1 + \min(\lambda \alpha, 1 - \alpha) \| \ell_R(0) \|_1. \] (27)
We will show that for \( \beta = 1 \), the right-hand side of (27) is strictly positive. This implies that the regulator’s threat is credible for \( \beta = 1 \). Moreover, since the left-hand side of (27) is continuous in \( \beta \), this implies that the threat is credible also for \( \beta < 1 \) sufficiently large.
For $\alpha \geq 1/(1+\lambda)$, the welfare loss $(1-\alpha)\ell^i(\xi)$ due to liquidation of bank $i$’s outside asset is smaller than or equal to $\lambda \xi^i$, hence the right-hand side of (27) is strictly positive as long as $\alpha < 1$. Suppose now that $\alpha < 1/(1+\lambda)$ and that (8) is satisfied instead. A quick calculation shows that (8) is equivalent to
\[(1-\alpha)\|e^{D\cup B}\|_1 < \lambda \alpha \|e^F\|_1 + \lambda \|V_0^{C\cup B}\|_1 - \lambda(1-\alpha)\|e^{C\cup B}\|_1.\] (28)

For any bank $i \in S$, define the function $f^i(x) = (1-\alpha)\ell^i(x) - \lambda x^i$ and observe that $f^i$ is decreasing until $x^i = V_0^i - e^i$ and increasing afterwards. Since $f^i(0) = 0$ for any $i \in S$ and $f^i(V_0 - (1-\alpha)e) > 0$ if and only if $i \in B$, this shows that $f^i(\zeta) \leq f^i(V_0 - (1-\alpha)e)$ for any $i \in B$ and any value of $\zeta^i$. We obtain the following auxiliary inequality
\[\lambda \zeta^i = (1-\alpha)\ell^i(\zeta) - f^i(\zeta) \geq (1-\alpha)\ell^i(\zeta) - f^i(V_0 - (1-\alpha)e), \quad i \in B, \] (29)
that we will use later. Next, we show that $(1-\alpha)\ell^i_N \leq \lambda \zeta^i$ for a bank $i \in S \setminus B$.
Indeed, if $\zeta^i \leq V_0^i - e^i$, then bank $i$ does not need to liquidate anything and hence $\ell^i_N = 0 \leq \lambda \zeta^i$ is trivially satisfied. For $\zeta^i > V_0^i - e^i$, the function $f^i$ is increasing and hence $f^i(\zeta) \leq f^i(V_0 - (1-\alpha)e) \leq 0$, where the latter inequality holds since $i \notin B$. We have thus shown that $(1-\alpha)\ell^i_N \leq \lambda \zeta^i$ for any $i \in S \setminus B$ and any $\zeta^i$. Together with (28) and (29), thus implies that
\[(1-\alpha)\|\ell_N\|_1 \leq (1-\alpha)\|e^{D\cup B}\|_1 + (1-\alpha)\|\ell^B_N\|_1 - (1-\alpha)\|\ell_B\|_1 - \lambda(1-\alpha)\|e^{C\cup B}\|_1 \leq \lambda \|e^{C\cup S\setminus B}\|_1 + \lambda \alpha \|e^F\|_1 + \sum_{i \in B} f^i(V_0 - (1-\alpha)e) \leq \lambda \|e^{C\cup S\setminus B}\|_1 + \lambda \alpha \|e^F\|_1.\]
Since $\|e^F\|_1 \leq \|\ell_R(0)\|_1$, this shows that the right-hand side of (27) is positive, thereby concluding the proof.

Proof of Lemma 4.5. Note first that
\[w_P = \lambda B + (1-\alpha)\|\ell_R(0)\|_1 - (1-(1+\lambda)\alpha)^+ \|\ell_R(0)\|_1.\]
Together with (10), this shows that
\[
\begin{align*}
  w_N - w_P &= (1 - \alpha) \| \ell_N - \ell_R(0) \|_1 + (1 - (1 + \lambda)\alpha)^+ \| \ell_R(0) \|_1 \\
  &\quad + \| \xi^{C_{ijs}} \|_1 - (1 + \lambda)(B - \| \delta^P_N \|_1).
\end{align*}
\]

For fixed $B$, the interbank repayments $\bar{p}_N$ are increasing in $\alpha$ and $\beta$ since a larger amount can be recovered and $\| \delta^P_N \|_1$ is decreasing in $\alpha$ and $\beta$ for the same reasons. Therefore, $\zeta^i = (\pi(L - \bar{p}_N))^i$ is decreasing and so is $\ell^i_N = \frac{1}{\alpha}(\zeta^i + e^i - V^i_0)^+$. Since $\ell^i_R(0)$ is constant for fixed $B$ and smaller or equal to $\ell^i_N$, the statement follows. \qed