NBER WORKING PAPER SERIES

OPTIMAL BAILOUTS AND THE DOOM LOOP WITH A FINANCIAL NETWORK

Agostino Capponi
Felix C. Corell
Joseph E. Stiglitz

Working Paper 27074
http://www.nber.org/papers/w27074

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2020

We would like to thank Árpád Ábrahám, Nicola Fuchs-Schündeln, Paul Glasserman, Piero Gottardi, Philipp Grübener, Philipp Kircher, Anton Korinek, Steve Kou, Lukas Nord, Giorgia Piacentino, our discussant Flora Lutz, conference participants at the 5th Annual Conference on Network Science in Economics, the SIAM Conference on Financial Mathematics & Engineering 2019, the 9th Workshop on Banks and Financial Markets (Vienna) and seminar participants at NYU and EUI for their helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Optimal Bailouts and the Doom Loop with a Financial Network
Agostino Capponi, Felix C. Corell, and Joseph E. Stiglitz
NBER Working Paper No. 27074
May 2020
JEL No. G01,G21,G28,H63,H81

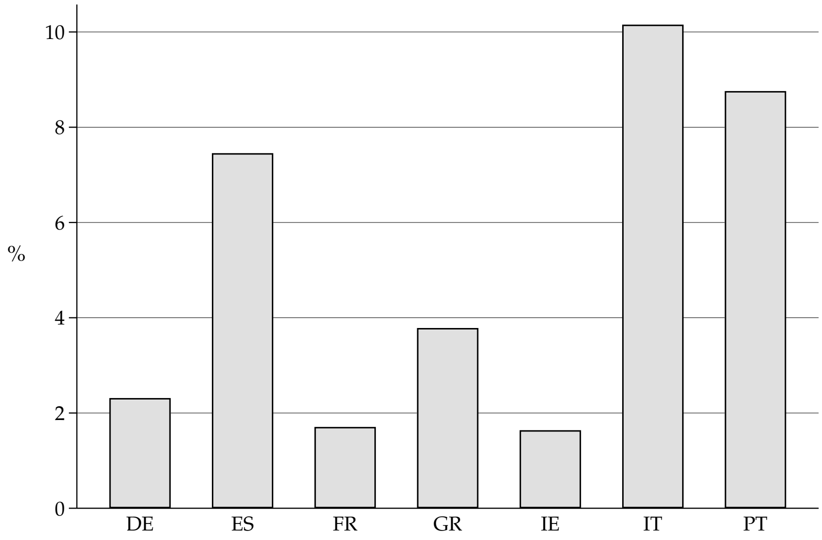
ABSTRACT

Banks usually hold large amounts of domestic public debt which makes them vulnerable to their own sovereign’s default risk. At the same time, governments often resort to costly public bailouts when their domestic banking sector is in trouble. We investigate how the interbank network structure and the distribution of sovereign debt holdings jointly affect the optimal bailout policy in the presence of this "doom loop". Rescuing banks with high domestic sovereign exposure is optimal if these banks are sufficiently central in the network, even though that requires larger bailout expenditures than rescuing low-exposure banks. Our findings imply that highly central banks can use exposure to their own government as a strategic tool to increase the likelihood of being bailed out. Our model thus illustrates how the "doom loop" exacerbates the "too interconnected to fail" problem in banking.

Agostino Capponi
Mudd Hall 535-G
Columbia University
500 W 120th St.
New York, NY 10027
ac3827@columbia.edu

Joseph E. Stiglitz
Uris Hall, Columbia University
3022 Broadway, Room 212
New York, NY 10027
and NBER
jes322@columbia.edu

Felix C. Corell
Department of Economics
European University Institute
felix.corell@eui.eu



Source: ECB Statistical Data Warehouse (June 2018)

Figure 1: Banks' domestic sovereign exposure (as a share of total assets)

1 Introduction

Throughout modern economic history, almost every major crisis has added its own signature terms to the dictionary of economic thought. For the European debt crisis that followed the Great Recession in 2008/09, one of the front-running candidates is certainly the so-called bank-sovereign "doom loop" - a term that has made its way into the vocabulary of policymakers, journalists, and researchers alike. The nexus between sovereigns and banks was identified as a major culprit responsible for the amplification of adverse shocks during the crisis.

The doom loop can be described as follows: Banks usually hold large amounts of domestic public debt which makes them vulnerable to sovereign default risk. For instance, as of June 2018 some eurozone banking sectors held well beyond 15% of their government's total stock of outstanding debt securities (see Figure 1). At the same time, governments often resort to costly public bailouts when their domestic banking sector is at stake. If these bailouts are debt-financed and lead to an increase in risk premia on sovereign debt (i.e., a decrease in bond prices), bank assets drop in value and the size of the bailouts required to keep the banks solvent increases further. If a significant fraction of banks' assets consists of domestic sovereign debt, this adverse feedback loop can leave both the government's and banks' balance sheets severely impaired. [1] Historical examples for the devastating force of the "doom loop" include Greece, where turmoil in sovereign debt markets damaged Greek

---

[1] Note for example that Italian banks currently hold around 10% of their assets in domestic sovereign debt (Figure 1).

banks' balance sheets and made large bailouts necessary that further weakened Greece's fiscal position; and Ireland, where the loop originated in the banking sector and thus worked in the opposite direction.

In this paper we investigate how the distribution of sovereign debt across banks and the structure of the interbank network jointly affect the government's optimal bailout decision in the presence of the doom loop. In doing so, we combine two strands of literature that have so far evolved separately, namely the doom loop literature (e.g. Acharya *et al.* (2014), Cooper and Nikolov (2018), Farhi and Tirole (2018)) and the literature about contagion in financial networks (Acemoglu *et al.* (2015), Elliott *et al.* (2014), Cabrales *et al.* (2017)). More precisely, we consider an exogenous network of (unsecured) interbank liabilities after an adverse shock has rendered a set of banks insolvent. Each bank holds an exogenous amount of its own government's bonds. Without any government intervention, there may be a default cascade that causes large welfare losses due to bankruptcy deadweight costs. The government can, however, reduce these losses by bailing out some or all of the failing banks. To do so, the government must borrow the required funds in the sovereign debt market which pushes down bond prices and thus puts further pressure on banks' balance sheets. A rational government internalizes this "second-round" effect, so bailouts are, by definition, large enough to cover the total shortfall of the banks to be rescued. The government trades off the welfare losses of a default cascade against those associated with a higher sovereign default probability, as effectively bank liabilities move onto the government's balance sheet.

Existing literature distinguishes two types of financial contagion, namely *direct* contagion due to counterparty default (see Glasserman and Young (2016) for a survey) and *indirect* contagion through common asset exposure (Adrian and Shin (2010); Greenwood *et al.* (2015)). Our paper features both types of contagion and emphasizes a link between both through the government's bailout decision. For example, a bailout may prevent direct contagion emanating from the targeted bank(s), but cause indirect contagion of banks that are critically exposed to the sovereign. On the other hand, our paper abstracts from more general macroeconomic effects (such as an increase in unemployment, a reduction in output, or a contraction of government revenues) that might emanate from any bailout policy. These effects were, of course, central to some of the earlier doom loop literature (see, e.g., Brunnermeier *et al.* (2016)).

Our first set of results applies to the special case of *complete* bailouts where the government can only bail out all banks or none. We show (not surprisingly) that the doom loop makes complete bailouts socially more costly if a larger amount of sovereign debt is held by domestic banks. Moreover, for a given aggregate amount of sovereign debt held by the banking sector, it is socially preferable that it be held by relatively well-capitalized banks. The intuition for these results is straightforward: Financing a bailout imposes a cost on banks that are exposed to sovereign default risk. The larger those banks' loss-absorbing equity buffer, the lower welfare losses will be. If, in contrast, public debt is largely held by weakly capitalized banks, the doom loop hits them with a stronger force and the required bailout may become very large.

Our second set of results generalizes the bailout decision by allowing the government to bail out any subset of banks (*partial* bailouts). The government chooses the set of surviving banks to maximize social welfare, taking into account impacts of the bailout of any bank on other banks and on sovereign risk. Our analysis shows that whether or not a given bank will be bailed out jointly depends on its position in the network and its sovereign debt exposure. We identify a measure of centrality ("node depth" as in Glasserman and Young (2015)) that captures a bank's contagion potential depending on its position in the network. Among two otherwise identical banks, the government prefers to bail out the one with higher node depth, even without of the doom loop. With the doom loop active, however, the presence of sovereign debt on banks' balance sheets - even if it is completely equally distributed - further strengthens centrality as a determinant of the government's bailout choice.

We also find that in general higher domestic sovereign exposure decreases banks' chances of being bailed out because of a "doom loop multiplier" effect: Every dollar raised for the bailout of a given bank lowers the value of sovereign debt on its balance sheet and thus increases the required bailout expenditure by a few cents. This effect is stronger for banks with larger domestic sovereign exposures. However, for systemically important banks there can be a countervailing force: Letting a systemically important bank fail *while others are bailed out* can lead to large additional deadweight losses among its creditors due to the associated drop in the value of its sovereign debt holdings. This finding implies that given a network of interbank liabilities, banks can use their sovereign debt position as a strategic tool to increase the odds of being bailed out. Our model thus provides a novel, network-based explanation for the increase in "home bias" in banks' sovereign portfolios during the European sovereign debt crisis.

2 Related Literature

Our work is related to the seminal articles by Kiyotaki and Moore (1997) and Allen and Gale (2000) in regards to the financial stability implications of different network structures. These studies use stylized networks to show how contagion propagates when individual financial institutions are hit by idiosyncratic negative liquidity shocks. Subsequent developments include Eisenberg and Noe (2001) who show existence and uniqueness of payment vectors that simultaneously clear liabilities in a general class of networks; Glasserman and Young (2015) who assess the extent of contagion accounting for bankruptcy costs; Acemoglu *et al.* (2015)

who focus on the systemic risk implications of different network topologies; and Elliott *et al.* (2014) and Cabrales *et al.* (2017) who model network linkages as equity cross-holdings or direct claims on other banks' projects. We refer to Cabrales *et al.* (2016) and Glasserman and Young (2016) for excellent surveys on contagion in financial networks.

Closely related to our paper is the small, yet growing, literature that analyzes government interventions in interbank networks. For example, Bernard *et al.* (2018) study under which conditions the government can credibly commit to organize an incentive-compatible *bail-in* scheme in which solvent banks contribute to rescuing the defaulting banks. Altinoglu and Stiglitz (2019) show how systemically important institutions can emerge as a result of banks expecting a public bailout. Erol (2019) shows that the expectation of bailouts leads to higher connectivity and a core-periphery network structure. In both of these papers the government's willingness to bail out the banks *ex post* makes the whole banking system more fragile *ex ante*. In contrast to our model, however, Erol (2019) does not model the government's bailout decision explicitly.

Our paper leverages insights from two strands of literature: the financial contagion literature and the large "doom loop" literature that has emerged in recent years in response to the European sovereign debt crisis. Brunnermeier *et al.* (2016) conceptually distinguish two types of loops. The first is the "real economy loop" where sovereign stress reduces the value of public debt on banks' balance sheets and thereby induces a credit crunch; through reduced economic activity and lower tax revenues, this credit crunch then feeds back into the fiscal position.[2] The second is the "bailout loop" that constitutes the focus of this paper.

Acharya *et al.* (2014) provide empirical evidence for both "directions" of the loop. They show that (a) bailouts triggered the rise of sovereign credit risk in 2008 and (b) changes in sovereign CDS rates in turn explain changes in bank CDS rates. The perceived stabilization of the banking sector through a bailout can thus turn out to be a "Pyrrhic victory".

Acharya *et al.* (2014) and Farhi and Tirole (2018) develop theoretical models of the "deadly embrace" that are related to ours, but do not consider interbank linkages and therefore the implications of network structure. Cooper and Nikolov (2018) study the strategic interaction between banks and the government. They show that if the government cannot commit to a credible no-bailout policy, banks in a subgame perfect Nash equilibrium anticipate a bailout and choose insufficient equity buffers. While in our model banks are inactive agents with exogenous balance sheets, our analysis of partial bailouts suggests that in a dynamic setting banks could influence the odds of being bailed out through their sovereign debt exposure.

Whereas Acharya *et al.* (2014), Farhi and Tirole (2018), and Cooper and Nikolov (2018) model the doom loop qualitatively in stylized three-period models, other studies have im-

---

[2] Indeed, Altavilla *et al.* (2017) and Popov and Van Horen (2015) find that during the European sovereign debt crisis banks with higher exposure to stressed sovereigns cut their lending to the domestic real economy significantly more than banks with low exposure.

plemented the feedback mechanism in full-blown quantitative macro models to evaluate the dynamic effects of banking regulation. In particular, Abad (2019) and Boz *et al.* (2014) conclude that higher capital requirements in the form of non-zero risk-weights for sovereign exposure (as stipulated by the Basel III framework) would improve welfare.

Gaballo and Zetlin-Jones (2016) and Thaler (2018) provide an interesting and novel perspective of the doom loop problem that is related to our model. They observe that higher domestic sovereign exposure of banks decreases the government's bailout capacity because the doom loop mechanism makes bailouts more costly. Interestingly, they even interpret high sovereign exposure as a commitment device that deters bank bailouts and thus alleviates moral hazard problems arising from bank behavior. Our benchmark results regarding *complete* bailouts can be interpreted similarly. However, our network approach allows us to go one step further and investigate what happens if some, but not all, banks may be bailed out. We show that in this scenario banks' incentives may be reversed: High sovereign exposure can *increase* an individual bank's chance of being bailed out, depending on its position in the interbank network.

Finally, our paper indirectly relates to a strand of literature that empirically tests different hypotheses of why euro area banks exhibited increasing sovereign debt "home bias" during and after the crisis. Altavilla *et al.* (2017) and Crosignani (2017) all find that the observed increase in domestic sovereign debt holdings during the crisis was stronger for poorly capitalized banks. This finding is worrisome if viewed in the context of our results on complete bailouts. As we show in the paper, bailouts are socially more costly if domestic sovereign debt is largely held by fragile banks.

3 Model

Our model has three periods $t = 0, 1, 2$ and features a set of banks $N = \{1, \ldots, n\}$ and a benevolent government that minimizes welfare losses after an exogenous shock has rendered a subset of banks insolvent.[3] The sequence of events in our model is illustrated in Figure 2. In the initial period $t = 0$, an exogenous negative shock hits a subset of the banks and lowers the value of their assets below that of their liabilities. Without a public bailout in $t = 1$ a cascade of bank defaults with associated bankruptcy deadweight costs may unfold. However, the government can decide to issue new debt in $t = 1$ and transfer funds to the insolvent banks to prevent the default cascade. In period $t = 2$ all sovereign debt matures and the government can either repay or not. We describe all relevant model components and the government's tradeoff in the following paragraphs.

---

[3] In that sense our model is closer to an Ireland-type doom loop than to the Greece-type, but we could just as well model the latter by initially shocking the government's fiscal position instead.

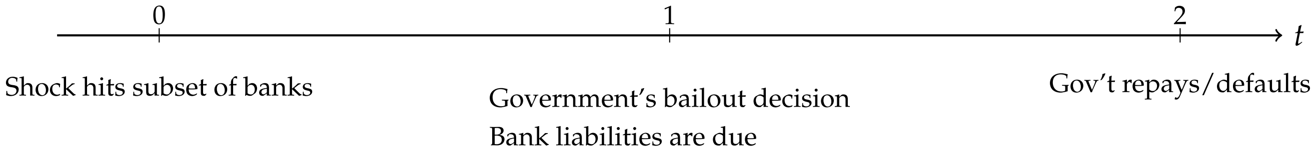


Figure 2: Timeline

3.1 Banks

The banks are connected through an exogenous network of unsecured interbank liabilities described by the matrix $\mathbb{L} = (L^{ij})$ for $i, j \in N$, where $L^{ij} \geq 0$ represents the gross value of bank j's liability to bank i and $L^{ii} = 0\ \forall i \in N$. We denote the total interbank liabilities of bank j by $L^j = \sum_{i=1}^n L^{ij}$ and use L to denote the corresponding $n \times 1$ vector. Moreover, it will be useful to define the relative liability matrix Π with entries $\pi^{ij} = L^{ij}/L^j$ if $L^j \neq 0$ and $\pi^{ij} = 0$ otherwise. In words, $\pi^{ij} \in [0, 1]$ is the share of bank j's interbank liabilities that it owes to bank i. Hence, the $n \times 1$ vector (ΠL) contains the book values of each bank's interbank assets. For instance, $(\Pi L)^i = \sum_{j=1}^n \pi^{ij} L^j$ is the book value of bank i's interbank assets.

In addition to interbank assets and liabilities bank i's balance sheet contains (a) outside assets $c^i \geq 0$ (including cash and loans to the non-financial sector), (b) senior liabilities $d^i > 0$ (e.g. deposits), and (c) exogenous domestic sovereign bond holdings $b^i \geq 0$ evaluated at the endogenous price q_t.[4] The value of equity at $t = 0$, defined as assets minus liabilities, is denoted by

$$V_0^i = \left(c^i + \sum_{j=1}^n \pi^{ij} L^j + q_0 b^i - d^i - L^i \right)^+ \quad \forall i \in N,$$

where we denote by $x^+ = \max(x, 0)$ the positive part of x. To simplify notation, we define $n \times 1$ vectors $d = (d^i)_{i=1}^n$, $c = (c^i)_{i=1}^n$, $b = (b^i)_{i=1}^n$ and $V_0 = (V_0^i)_{i=1}^n$. A stylized balance sheet at time $t = 0$ is shown in Figure 3.

A negative shock (e.g. a natural disaster) has reduced the value of banks' outside assets to c and thus rendered some of them insolvent, i.e., their equity is wiped out and they face a shortfall of $\chi_0 \equiv (L + d - c - q_0 b - (\Pi L))^+$. Since all bank liabilities are due at $t = 1$ the insolvent banks will have to default on their obligations if they are not bailed out. We

---

[4] In practice, whether or not securities are "marked to market" (MTM) depends on whether they are categorized as "held to maturity" (HTM, evaluated at amortized cost) or "available for sale/trading" (evaluated at current market prices). In our model banks may have to sell sovereign debt to settle their liabilities in $t = 1$, so MTM is the appropriate concept.

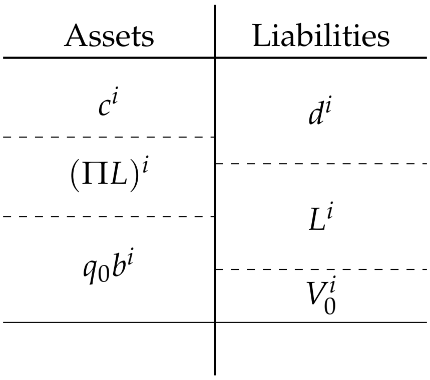


Figure 3: A *solvent* bank i's balance sheet at t=0

denote the set of these *fundamentally* defaulting banks by $\mathcal{F} \equiv \{i | \chi_0^i > 0\}$ and maintain the assumption that $\mathcal{F}$ is non-empty throughout the paper.

Default Cascades

Since banks are interlinked through credit contracts, the default of one or more fundamentally defaulting banks can lead to further defaults elsewhere in the system. If bank j defaults, it has to liquidate all its (interbank and other) assets and repay its creditors according to their seniority. The financial commitments d^j have seniority over interbank liabilities L^j. Among interbank creditors, the total repayment $p^j < L^j$ is distributed *pro rata*, i.e. the creditor bank i of a defaulting bank j is repaid according to its share π^{ij} in bank j's interbank liabilities.

As in Glasserman and Young (2015) we assume that the bankruptcy of bank j comes at a deadweight loss of $\beta\chi_1^j$ where $\chi_1^j \equiv (L^j + d^j - c^j - q_1 b^j - \sum_{k=1}^n \pi^{jk} p^k)^+$ is bank j's shortfall in period $t = 1$ and $\beta > 0$ a scaling parameter. The deadweight loss is increasing in the bank's shortfall and can be interpreted as legal costs or losses due to an inefficient allocation of resources during the bankruptcy procedure.[5] Of course, the deadweight loss caused by a defaulting bank is bounded from above by the total value of its assets. To make comparative statics tractable, we focus on settings where this upper bound is not binding, i.e., the initial shortfalls χ_0^i and the scaling parameter β are small enough to ensure that no bank's assets are completely wiped out in bankruptcy.

We assume that interbank liabilities are cleared simultaneously at $t = 1$, as in Eisenberg and Noe (2001). This leads to the following definition:

---

[5] See Bernstein *et al.* (2019) for a recent empirical study.

Definition 1. A *clearing payment vector* $p = (p^1, \dots, p^n)$ for a financial system (L, Π, c, d, b) is a fixed point of

$$p^i = \begin{cases} L^i & \text{if} \quad c^i + q_1 b^i + \sum_{j=1}^n \pi^{ij} p^j \geq L^i + d^i \\ \left(c^i + q_1 b^i + \sum\limits_{j=1}^n \pi^{ij} p^j - d^i - \beta \chi_1^i \right)^+ & \text{otherwise} \end{cases}$$

The existence of multiple, Pareto ranked clearing payment vectors in this setting follows directly from Proposition 3 in Glasserman and Young (2015). Following the standard convention in this literature, we focus on the Pareto dominant clearing payment $\bar{p}$, i.e., the "best case" scenario with the highest repayments. We would, however, like to emphasize that this represents only a lower bound on losses suffered by banks in a default cascade. In reality the uncertainty about which clearing payment vector will realize can be a great concern for policymakers and tilt the government's decision towards a bailout.

If the recovery value of bank i's assets is lower than its senior obligations d^i, the junior creditors receive nothing ($\bar{p}^i = 0$) and senior creditors suffer a loss of

$$\delta^i(\bar{p}) = \left(d^i - \left(c^i + q_1 b^i + \sum_{j=1}^n \pi^{ij} \bar{p}^j - \beta \chi_1^i \right)^+ \right)^+ . \tag{1}$$

We define the set of defaulting banks under a clearing payment vector $\bar{p}$ as $\mathcal{D}(\bar{p}) \equiv \{i | \bar{p}^i < L^i\}$. Clearly, $\mathcal{D}(\bar{p}) \supseteq \mathcal{F}$ because $\mathcal{D}(\bar{p})$ can also include banks that default due to contagion effects triggered by fundamentally defaulting banks.

3.2 Government

The government in our model faces a tradeoff in minimizing expected welfare losses at $t = 1$. On the one hand, it wants to avoid bank defaults and the associated deadweight losses. On the other hand, rescuing banks with debt-financed bailouts is also costly because liabilities are effectively moved from private to the public balance sheet, resulting in higher sovereign spreads and a higher probability of a costly sovereign default.

We assume that all sovereign debt matures in the final period $t = 2$ and is held either by domestic banks or risk-neutral investors (external to the network), so the (exogenous) total initial stock of sovereign debt is given by $B_0 = \sum_{i=1}^n b^i + b_0^{Inv}$, where b_0^{Inv} denotes the holdings of risk-neutral investors.

At $t = 1$, after observing the fundamental shocks to the banking system, the government can choose to prevent a default cascade by issuing additional debt, i.e., by increasing the outstanding debt level to $B_1 > B_0$ to finance bank bailout transfers $t^i(q_1) \geq 0$ for $i \in N$.

In the first part of the paper we focus on *complete* bailouts, i.e., we restrict the government's set of possible bailouts to those where *all* banks receive exactly the amount required to cover their shortfall:

$$t^i(q_1) = \left(L^i + d^i - c^i - q_1 b^i - (\Pi L)^i\right)^+ \quad \forall i \in N. \tag{2}$$

This form of bailout is extreme (saving either all banks or none), but constitutes a useful benchmark for comparing our model with standard representative bank models in which the bailout decision is often binary. What is more, in reality discriminating among banks might lead to litigation that governments seek to avoid. Later in the paper, however, we extend the analysis to the more general case where the government can optimally choose the set of surviving banks. Notice that the set of banks receiving a positive bailout transfer is not necessarily limited to $\mathcal{F}$. It is evident from equation (2) that if the price of sovereign debt falls in period 1 ($q_1 < q_0$), even banks which do not belong to the set $\mathcal{F}$ may have to be rescued to avoid a default cascade. To raise the total amount of required bailout $T(q_1) \equiv \sum_{i \in N} t^i(q_1)$, public borrowing needs to satisfy the following budget constraint:

$$q_1(B_1 - B_0) = T(q_1) \tag{3}$$

Clearly, the amount of debt that the government can raise by issuing $(B_1 - B_0)$ new bonds in period 1 depends on q_1, i.e., the (endogenous) market price of debt prevailing in period 1. In the remainder of this section, we first explain how sovereign bonds are priced and then how the doom loop emerges in our model.

In the final period $t = 2$ the government raises taxes to repay its obligations B_1. We assume that the government's tax capacity (i.e., the maximum revenue it can raise) $\tilde{\tau}$ is a random variable that follows a continuous and strictly monotonic cumulative distribution function (CDF) $F(\tau)$.[6] If $\tau \geq B_1$, the government raises exactly enough to repay all debt in full. If, however, $\tau < B_1$, we assume the government collects nothing and *fully* defaults on all its obligations, i.e. bondholders do not receive any repayment at all.[7] This is consistent with the fact that in the aftermath of a sovereign default and the ensuing crisis, the ability of the government to function is greatly circumscribed. It then follows that the probability of a sovereign default is given by $P(Default) = F(B_1) = F\left(B_0 + \frac{T(q_1)}{q_1}\right)$. Because the cumu-

---

[6] One interpretation would be that the government can raise any revenue it desires, but only at a stochastic cost. In this alternative scenario the government would optimally default whenever the realization of this cost exceeds some threshold.

[7] This assumption buys us tractability and leads to a simple analytical expression for q_1. If one were to assume partial repayment, the numerator in equation (4) would also include the integral over all possible realizations of $\tilde{\tau}$ and the corresponding pro rata repayments. The economic mechanism we want to capture is qualitatively unaffected by this "all-or-nothing" assumption on tax collection.

lative distribution function is increasing, it follows immediately that the default probability increases with the required bailout expenditures $T(q_1)$, but decreases with respect to the bond price q_1. In words, a higher bond price in period $t = 1$ will make a sovereign default less likely. At the same time, however, the bond price depends on the government's default probability in the following way: The marginal buyers of sovereign bonds are investors who discount future cash flows at the risk-free gross interest rate $R \geq 1$. Because we assume that these investors are risk-neutral and have deep pockets, the market-clearing sovereign debt price is given by

$$q_1 = \frac{1 - P(Default)}{R} = \frac{1 - F\left(B_0 + \frac{T(q_1)}{q_1}\right)}{R}. \tag{4}$$

At this price q_1 risk-neutral investors are indifferent between buying the bond with expected payoff $1 - P(Default)$ and the risk-free asset that guarantees the return R.

Equation (4) is a nonlinear equation which generally admits multiple solutions. As a result, there may exist multiple equilibrium sovereign debt prices, just like in the canonical models of self-fulfilling debt crises such as Calvo (1988) or Cole and Kehoe (2000). This multiplicity arises from the fact that both sides of equation (4) are increasing in q_1. In economic terms, if investors expect a low default probability they are willing to pay a high price q_1 for sovereign bonds. According to the right-hand side of equation (4), this confirms the expectation of a low default probability. The reverse argument holds when investors believe in a high default probability.[8]

How many equilibrium prices exist depends on the exogenous parameters R and B_0, the bailout transfer in (2) and, importantly, the shape and domain of $F(\tau)$. We make the following assumption:

Assumption 1. *For a given network* (L, Π, c, d, b) *and parameters* (B_0, R)*, the tax capacity* $\tilde{\tau}$ *follows a Pareto distribution with a heavy tail, i.e.,*

$$F(\tau) = 1 - \left(\frac{B_0}{\tau}\right)^{\alpha}$$

with $\tau \geq B_0$ *and* $\alpha < 1$*.*

This assumption has several implications. First of all, it imposes a lower bound on the space of possible realizations of $\tilde{\tau}$ and implies $F(B_0) = 0$, so it ensures that without an increase in debt the government can always repay its debt with certainty. As a consequence,

---

[8] For a similar analysis of the effect of nominal interest rates on the price of private bonds, see Greenwald and Stiglitz (1993).

without bailouts the equilibrium price in the sovereign debt market is pinned down at $q_1^{NB} = R^{-1}$, i.e., the government pays no spread relative to the risk-free rate R.[9]

Second, because $F(B_0 + \varepsilon) > 0 \ \forall \varepsilon > 0$, if public debt increases (e.g. due to a positive bailout at $t = 1$) the government will default at $t = 2$ with a non-zero probability. In other words, even a small bailout always comes at the cost of increasing sovereign spreads.

Finally, the parameter $\alpha < 1$ guarantees that the right-hand side of equation (4), when viewed as a function of q_1, has the graph depicted in Figure 4. In particular, the curve starts above the 45° line and is strictly concave. Assumption 1 thus implies a unique equilibrium price with bailouts $q_1^B \in (0, R^{-1})$ which is decreasing in the level of bailout expenditures, i.e., a higher price q_1^B can only be achieved through lower bailout expenditures and vice versa.[10]

Lemma 1. *Equilibrium aggregate bailout transfers $T(q_1)$ and the equilibrium price of sovereign debt q_1^B are inversely related.*

Proof. See Appendix B. ■

Note that we chose the Pareto distribution for analytical tractability, but any other power law distribution with sufficiently slow decay would also yield a unique equilibrium price under mild extra conditions. Even with more general distributions that would generate multiple (stable or unstable) equilibrium prices, our results would still survive locally for the doom loop equilibria with $q_1 < R^{-1}$. For an extensive discussion, see Appendix A.

In economic terms the Pareto assumption means that there is a high probability of being able to raise small extra tax revenues (beyond a guaranteed minimum B_0), but that higher revenues become less and less likely. The heavy tail of the Pareto distribution implies that even very high tax revenues are not completely unlikely, so the government might be able to honor extremely large debt obligations.

Graphical Illustration of the Doom Loop

Figure 4 illustrates the equilibrium in the sovereign debt market and the result from the lemma graphically for different bailout amounts. For positive bailout $\tilde{T}$ (solid line), a locally stable equilibrium price $q_1^B = \tilde{q}_1 < 1/R$ emerges. Most importantly, note that higher bailout expenditures $\hat{T} > \tilde{T}$ shift the curve downwards (dashed line) and reduce the equilibrium

---

[9] This normalization is without loss of generality. When the government considers a bailout it is concerned with the implied change in its default probability, not with the level. A positive default probability without bailouts (e.g. because bank failures lead to a macroeconomic contraction with lower tax revenues and higher government expenditure) would simply add a level shift to the welfare losses without bailout which will be defined in (6).

[10] Strictly speaking, a "market breakdown" equilibrium in which $q_1^B = 0$ (the government defaults with certainty) always exists for $T > 0$ as $\lim_{\tau \to \infty} F(\tau) = 1$ because $F(\cdot)$ is a CDF. In our analysis we deliberately ignore this extreme coordination failure.

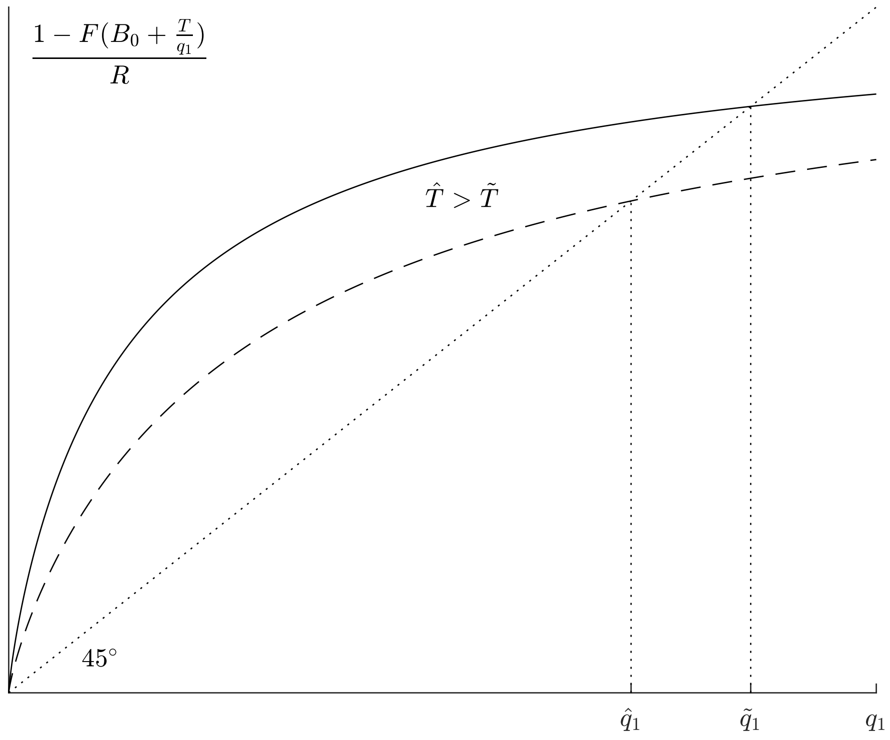


Figure 4: Larger bailout expenditures reduce sovereign debt price

debt price to $\hat{q}_1 < \tilde{q}_1$. Because from equation (2) the required bailout amount is inversely related to q_1, the figure illustrates how the doom loop operates in the model: Bailouts lead to higher spreads ($q_1 \downarrow$) which in turn increases the required bailout expenditures. This pushes down q_1 further and so forth.

What remains to be defined is the initial sovereign debt price q_0. We assume that the shock to the banking sector (and thus a potential bank bailout) is totally unexpected, so it is not reflected in the pre-shock price of sovereign debt. Therefore, the initial bond price is given by

$$q_0 = \frac{1 - F(B_0)}{R} = \frac{1}{R}. \tag{5}$$

As a consequence, whenever $T(q_1) > 0$, the sovereign spread jumps up in $t = 1$. In contrast, whenever $T(q_1) = 0$, we have $q_1^{NB} = q_0 = 1/R$.

Welfare Losses

The government chooses the bailout strategy in period $t = 1$ to minimize welfare losses. In the absence of intervention, welfare losses are given by

$$w(\bar{p}) = \beta \sum_{i \in \mathcal{D}(\bar{p})} \left(L^i + d^i - c^i - q_1 b^i - \sum_{j=1}^{n} \pi^{ij} \bar{p}^j \right) = \beta \sum_{i \in \mathcal{D}(\bar{p})} \chi_1^i, \tag{6}$$

the part of defaulting banks' assets that is lost in bankruptcy. Note that we could also allow bank failures to affect the public balance sheet, e.g. through a macroeconomic contraction and an increase in automatic stabilizer expenditures. This would increase $w(\bar{p})$ and thus tilt the government's incentives further toward a bailout.

If the government *does* bail out all banks, liabilities are effectively shifted from the banking sector to the public, there are no bankruptcies and (expected) welfare losses are instead given by

$$w^B = \gamma B_1 P(Default) = \gamma \left(B_0 + \frac{T(q_1^B)}{q_1^B} \right) F \left(B_0 + \frac{T(q_1^B)}{q_1^B} \right), \tag{7}$$

where the parameter $\gamma > 0$ measures deadweight losses per unit of defaulted sovereign debt and is a proxy for litigation costs and the penalty of losing access to debt markets.[11] Therefore, according to (7) high spreads (low q_1^B) are costly in our model not because they constrain government expenditures, but because they (a) make a sovereign default more likely and (b) increase the cost of a sovereign default (since the government will have had to issue more bonds to finance the bailout). The government internalizes the impact of its bailout choice on the default probability and finances a bailout if and only if $w^B < w(\bar{p})$.

An alternative version could feature an additional "deposit insurance" γ-term in equation (6), namely if depositor losses from (1) automatically triggered public transfers that raise public debt beyond B_0. Qualitatively, the only difference would be that in such an alternative specification the government would focus its bailout efforts more strongly on banks with otherwise large depositor losses.

Similarly, one could introduce loans to the real economy as an additional bank asset that the government seeks to sustain in a crisis. In that case, banks that are crucial for credit provision would likely receive a bailout.[12]

A third alternative would highlight the fact that a sovereign default might come with additional benefits: There is a reduction in the need to raise taxes to service the debt. Our parameter γ can be interpreted as the *net* deadweight loss associated with a sovereign default, after taking into account the benefit of reduced taxation.

3.3 Equilibrium with Complete Bailouts

We start by analyzing the situation where the government is restricted to choose between no bailout or a complete bailout. The corresponding definition of equilibrium is given next.

---

[11] It is reasonable to assume that the cost of a sovereign default increases with the level of defaulted debt, but the linear functional form in (7) is not crucial. Our results would also hold for any non-decreasing default cost function $\gamma(B_1)$.

[12] Since large corporations can access capital markets more easily or benefit from central bank support for the commercial paper segment, such a specification would put greater emphasis on small and regional banks that supply loans to small and medium enterprises.

Table 1: How does the bailout space Ψ depend on model parameters?

α	γ	β	R
-	-	+	+/-

Note: The signs denote whether the bailout space increases (+) or decreases (-) with the respective parameter. Parentheses indicate weak monotonicity. Ambiguous effects are denoted by +/-. Whenever we say that Ψ increases with a parameter θ it means that for $\theta'' > \theta'$ we have $\Psi'' \supset \Psi'$ and vice versa.

Definition 2. For a given network (L, Π, c, d, b), bankruptcy cost parameter β, risk-free interest rate R, and fiscal parameters (B_0, α, γ), an *equilibrium with complete bailouts* is a set of bailout transfers $t^i(q_1) \geq 0$ for $i \in N$ and sovereign debt prices q_0 and q_1 such that

1. Bailouts $\left(t^i(q_1)\right)_{i \in N}$ satisfy equation (2) if $w^B < w(\bar{p})$ and $t^i = 0 \; \forall i \in N$ otherwise;
2. Sovereign debt prices are determined by equations (4) and (5).

Before we turn to the analysis of how the sovereign debt distribution affects the strength of the doom loop and thus the government's bailout decision, it is helpful to define the *bailout space*, i.e. the set of initial shocks for which the government will bail out the banks in equilibrium. We define the bailout space as a subset of the space of vectors c (i.e., "cash" assets) for a given network structure and parameters. Everything else equal, a (component-wise) weakly smaller vector c corresponds to a larger initial shock.

Definition 3. For given (L, Π, b, d), bankruptcy cost parameter β, risk-free interest rate R, and fiscal parameters (B_0, α, γ), the *bailout space* is defined as

$$\Psi \equiv \left\{ c \in \mathbb{R}^n : w^B < w(\bar{p}) \right\}$$

Note that the topological properties of the bailout space are not obvious. Specifically, a stronger initial shock $c'' < c' \in \Psi$ is ambiguous: On the one hand, it will create (weakly) more damage in a default cascade ($w(\bar{p}) \uparrow$) which suggests that c'' might still belong to the bailout space. On the other hand, it also (weakly) increases the fiscal cost of a bailout, suggesting the opposite ($w^B \uparrow$ and $c'' \notin \Psi$). Nevertheless, the bailout space depends on some model parameters in a straightforward way. We summarize these comparative statics results in Table 1, but relegate the corresponding proofs to Appendix B.

An increase in the Pareto shape parameter α induces a first-order stochastic dominated shift of the tax capacity distribution $F(\tau)$, i.e., higher taxation capacities become less likely. As a consequence, the probability of sovereign default and hence w^B increase, so bailouts become less attractive and Ψ decreases.

A higher γ makes an eventual sovereign default more costly ($w^B \uparrow$), so again Ψ shrinks as γ increases. In contrast, a higher bankruptcy cost parameter β means that a potential default cascade results in higher welfare losses ($w(\bar{p})$ ↑) which makes bailouts relatively more attractive and increases Ψ. The European Union's recent efforts to create a banking union with a robust resolution framework can partially be understood as an attempt to bring down β and thereby to discourage public bailouts in the first place.

Finally, the effect of R on the bailout space is ambiguous. A higher risk-free gross interest rate R increases both $w(\bar{p})$ (because it lowers the value of banks' sovereign debt holdings) and w^B (because of a higher default probability and larger deadweight losses in case of default).

The Doom Loop with Complete Bailouts

Our main research question is how the interbank network topology and the distribution of sovereign debt jointly affect the government's bailout choice. In the context of complete bailouts, answering this question boils down to evaluating the relative effects of sovereign debt distribution and network structure on $w(\bar{p})$ and w^B.

First note that in the case of complete bailouts all interbank liabilities are paid in full, so the network structure only matters for $w(\bar{p})$, as the required bailout that determines w^B is independent of the network structure. In contrast, the presence and distribution of sovereign debt only affects w^B as without a bailout there is no doom loop and we have $q_1^{NB} = 1/R$ in equilibrium. Specifically, what matters for w^B is (a) how much debt the banking sector holds, and (b) which banks hold the debt. We address both points in the following two propositions.

Proposition 1 (Sovereign exposure and the bailout space). *For a constant initial level of public debt B_0, let $(L, \Pi, \tilde{c}, d, \tilde{b})$ and $(L, \Pi, \hat{c}, d, \hat{b})$ be two financial systems such that*

1. $\hat{b}^i \geq \tilde{b}^i \quad \forall i \in N$ *with at least one inequality strict*
2. $\hat{c}^i = \tilde{c}^i - (\hat{b}^i - \tilde{b}^i)/R > 0 \quad \forall i \in N : \hat{b}^i > \tilde{b}^i$.

Then with complete bailouts $\hat{w}^B \geq \tilde{w}^B$ and $\hat{\Psi} \subseteq \tilde{\Psi}$.

Proof. See Appendix B ■

Conditions 1 and 2 ensure that if we move from the tilde-system to the hat-system, every bank that now holds more sovereign debt holds less cash in return so that the overall initial equity or shortfall positions remain unchanged ($\hat{\chi}_0 = \tilde{\chi}_0$). In other words, these banks only differ in their asset composition across the two scenarios, but not in their total asset size or leverage. The idea of the proposition is that increasing the amount of sovereign debt held by domestic banks amplifies the feedback loop between a falling debt price q_1 and ever higher transfers $T(q_1)$. Therefore, complete bailouts become more costly and as a result Ψ becomes smaller.[13] Note however that $\hat{w}^B$ is only weakly larger than $\tilde{w}^B$ because it is possible that the banks with increased sovereign exposure are well capitalized and do not require a bailout in the first place. In that case, even though their shareholders suffer higher book value losses than in the tilde-system, these do not translate into outright bankruptcies and hence do not affect the bailout space.

Clearly, if the banks with sovereign exposure are relatively well capitalized, they can better absorb a drop in q_1 than weak banks. This higher loss-absorbing capacity mitigates the "doom loop"-related cost of a bailout because fewer banks need to be bailed out (see equation (2)). Even if bailing out the fundamentally defaulting banks makes additional bailouts necessary (i.e. $t^i(q_1) > 0$ for some $i \notin \mathcal{F}$ due to $q_1 \downarrow$), this bailout will still be less costly than if the sovereign debt was held entirely by the $\mathcal{F}$-banks. We present an example to illustrate the economic forces at play.

Example 1. *Let $\mathcal{B} \equiv \{i : b^i > 0\}$ denote the set of banks that hold sovereign debt and consider the following scenario: Suppose there was only one fundamentally defaulting bank i while all sovereign debt in the banking sector was held by another bank $j \neq i$, so $\mathcal{F}$ and $\mathcal{B}$ would be disjoint singletons. Suppose further that the government chooses to bail out bank i, thereby causing a drop in the bond price that would render bank j insolvent. Then, according to our definition of complete bailouts, the government also has to transfer bailout funds to bank $j \in \mathcal{B}$. However, as long as bank j initially had a positive equity buffer, the total required bailout $(t^i + t^j)$ will be smaller than if only bank i had held the sovereign debt stock (i.e. if $\mathcal{F} = \mathcal{B} = \{i\}$) and the doom loop had hit bank i with its full force.*

The insight from this stylized example carries over to more dispersed distributions of sovereign debt among banks. We formally state this intuitive result in the following proposition:

Proposition 2 (Sovereign exposure and capital buffers)**.** *Let $(L, \Pi, \tilde{c}, d, \tilde{b})$ and $(L, \Pi, \hat{c}, d, \hat{b})$ be two financial systems such that*

1. $\hat{b}^i < \tilde{b}^i \; \forall i : V_0^i < \tilde{b}^i(R^{-1} - \tilde{q}_1^{max})$
2. $\sum_{i \in N} \hat{b}^i = \sum_{i \in N} \tilde{b}^i$
3. $\hat{c}^i = \tilde{c}^i - (\hat{b}^i - \tilde{b}^i)/R > 0 \quad \forall i \in N : \hat{b}^i < \tilde{b}^i,$

where $\tilde{q}_1^{max}$ is the unique positive solution of $R\tilde{q}_1^{max} = 1 - F\left(\frac{\sum_{i \in \mathcal{F}} \chi_0^i}{\tilde{q}_1^{max}}\right)$. Then with complete bailouts $\hat{w}^B < \tilde{w}^B$ and $\hat{\Psi} \supset \tilde{\Psi}$.

Proof. See Appendix B ■

---

[13] This relationship perfectly resonates with the argument in Gaballo and Zetlin-Jones (2016) that stronger "home bias" in banks' sovereign debt portfolios reduces the government's bailout capacity.

The two financial systems described in the proposition only differ in the distribution of domestic sovereign debt, with aggregate sovereign exposure in the banking system and initial shortfalls held constant (conditions 2 and 3 imply $\hat{\chi}_0 = \tilde{\chi}_0 \equiv \chi_0$). More precisely, the banks with capital buffers V_0^i below an exogenous threshold hold less sovereign debt in the hat-system than in the tilde-system (condition 1), which requires that banks with V_0^i above the threshold hold more.[14] Then, the proposition says that redistributing domestic sovereign debt from weakly capitalized banks to more healthy banks makes a complete bailout socially less costly. The intuition behind this result is that *if banks with sovereign exposure have larger capital buffers they can absorb any drop in the value of sovereign debt more easily without becoming insolvent themselves (and thus requiring a bailout).* Therefore, aggregate bailouts and welfare losses are smaller. Proposition 2 can be thought of in terms of the correlation between V_0^i and b^i: Everything else equal, higher correlation leads to less costly bailouts.

We conclude our discussion of the benchmark case with complete bailouts by briefly touching upon the policy dimension of our findings. Since we consider an ex post scenario with a given distribution of sovereign debt, the correlation between the elements of V_0 and b is just a primitive of our model. However, proposition 2 suggests that from an *ex ante* perspective (with uncertainty about idiosyncratic fundamental shocks) it is socially preferable that domestic sovereign debt be held by the "safest" banks, i.e. those that are the least likely to default fundamentally. Against this backdrop, the findings in Acharya and Steffen (2015), Altavilla *et al.* (2017), and Crosignani (2017) that in the recent European sovereign debt crisis the most fragile banks increased their domestic sovereign exposure disproportionally becomes even more worrisome; rather than "too big to fail" they might have become "too debt-loaded to be saved". Our results therefore provide an immediate rationale for bank capital requirements that increase with domestic sovereign debt exposure. We refer to Véron (2017) for a detailed proposal and a comprehensive survey of the policy debate.

3.4 Optimal (Partial) Bailouts

In the previous section we have only considered binary outcomes: The government either chooses to rescue all banks or none. As the global financial crisis of 2007-2008 has shown in several countries, reality may lie somewhere between those two extremes. In some situations, complete bailouts can be unrealistically costly in our model, so they often do not occur in equilibrium. Therefore, in this section we expand the set of feasible bailout strategies be-

---

[14] This threshold depends on $\tilde{q}_1^{max}$, i.e., the upper bound on the bond price in case of a bailout. This is the price that would prevail if only the initial shortfall had to be covered because none of the banks in $\mathcal{F}$ held any sovereign debt. Hence, the banks referred to in condition 1 are those whose capital buffers could not even absorb the lowest possible drop in the sovereign bond price.

yond complete bailouts and allow the government to pick the optimal subset of "surviving" banks $\mathcal{S} \subseteq N$.

As in the case of complete bailouts, welfare is maximized by the smallest public transfer that makes the surviving banks solvent. Hence, the formula for bailout transfers takes a form similar to equation (2). However, note that in contrast to the complete bailouts benchmark there may now be defaulting banks (i.e. the complement of $\mathcal{S}$ which we denote by $\mathcal{D}(\bar{p}(\mathcal{S}))$) that do not pay their interbank liabilities in full. Partial bailout transfers account for these losses on interbank assets:

$$t^i(\bar{p}(\mathcal{S}), q_1) = \left(L^i + d^i - c^i - q_1 b^i - (\Pi \bar{p}(\mathcal{S}))^i\right)^+ \quad \forall i \in \mathcal{S} \tag{8}$$

Even though partial bailouts defined in this way are a generalization of complete bailouts (obtained by setting $\mathcal{S} = N$), the government cannot choose the size and distribution of bailout transfers arbitrarily. In particular, we do not allow transfers to banks outside of the set $\mathcal{S}$, i.e., the government can only provide bailout funds to banks that become solvent after these transfers.[15] This modeling choice can be interpreted as a political economy constraint: Voters would not allow their elected officials to acquire a stake in a bank that is certain to fail anyway.

The government will choose the set $\mathcal{S}$ (and equivalently its complement $\mathcal{S}^c = N \setminus \mathcal{S}$) to minimize the expected welfare losses $w(\bar{p}(\mathcal{S}))$ arising from bank bankruptcies and sovereign default. Formally, the government problem becomes

$$\begin{aligned} \min_{\mathcal{S} \subseteq N} \quad & \beta \sum_{i \in \mathcal{S}^c} \left(L^i + d^i - c^i - q_1 b^i - \sum_{j=1}^{n} \pi^{ij} \bar{p}^j(\mathcal{S}) \right) + \gamma B_1 P(Default) \\ s.t. \quad & \mathcal{S} = \mathcal{D}(\bar{p}(\mathcal{S}))^c, \end{aligned} \tag{9}$$

where the constraint needs to be imposed for the sake of technical consistency because not every subset of N can solely survive.[16] We are now in a position to define equilibria with optimal bailouts:

Definition 4. For a given network (L, Π, c, d, b), bankruptcy cost parameter β, risk-free interest rate R, and fiscal parameters (B_0, α, γ), an *equilibrium with optimal bailouts* is a set of surviving banks $\mathcal{S}^*$, associated bailout transfers $t^i(\bar{p}(\mathcal{S}^*), q_1) \geq 0$ for $i \in \mathcal{S}^*$, and sovereign debt prices q_0 and q_1 such that

1. $\mathcal{S}^*$ solves the government's problem (9);

---

[15] At the optimal $\mathcal{S}^*$, if the marginal benefit of bailout funds in an insolvent bank exceeds the marginal cost of raising these funds, removing the constraint would actually allow for a welfare improvement. However, it would also make the problem much less tractable.

[16] Some banks may never fail (so they automatically have to be included in each candidate $\mathcal{S}$) or some always survive together so that it is impossible for one to be in $\mathcal{S}$ while the other is not.

2. Bailouts $\left(t^i(\bar{p}(\mathcal{S}^*), q_1)\right)_{i\in\mathcal{S}^*}$ satisfy equation (8);
3. Sovereign debt prices are determined by equations (4) and (5).

Local Analysis of Bailout Determinants

The optimization problem in (9) is a highly complex minimization over sets that involves several nonlinearities and fixed points. Hence, it is generally difficult to make general analytical statements about the determinants of $\mathcal{S}^*$.[17]

In particular, it is hard to pin down whether or not a given bank will be bailed out in equilibrium. Therefore, to develop a better understanding of the economic forces that determine the government's bailout decision, from now on we ignore the "global" solution of (9) and focus instead on "local" analysis in the following sense: Suppose that instead of choosing the optimal subset among all feasible subsets of surviving banks, the government could only bail out one bank at a time. Starting from the hypothetical laissez faire outcome with defaulting banks $\mathcal{D}(\bar{p})$, the government computes for every troubled bank the net welfare gain that would result from bailing it out. As long as a reduction in welfare losses is possible in this way, the government adds the bank with the highest net welfare gain to the set of surviving banks. The procedure stops once there is no troubled bank left that can be bailed out without increasing welfare losses.

There is no guarantee that the set thus constructed coincides with the "globally" optimal set $\mathcal{S}^*$ which solves (9).[18] However, the fictional sequential procedure allows us to at least analyze "local" optima, i.e., we can pin down bank-specific properties that determine which bank will be bailed out at any given step of the procedure.

To make this analysis meaningful, in the following we only consider situations where *partial* bailouts are optimal, i.e., some banks are bailed out in equilibrium while others are not. It is easy to see from (9) that high values of β relative to γ make complete bailouts attractive whereas low values will prevent bailouts altogether. Figure 5 plots the regions of β and γ for which the equilibrium outcome is a complete bailout, a partial bailout, or no bailout, for a numerical example. We will focus on the intermediate region in the γ-β plane, that is the light grey area in the figure.

We begin the analysis with a useful decomposition of the *net welfare effect* of bailing out bank i which we call $\Delta^i w$.[19] On the one hand, the bailout brings a benefit by not only avoiding direct bankruptcy deadweight losses $\beta\chi_1^i$, but also by increasing bank i's interbank re-

---

[17] Of course, it is possible to solve (9) numerically. For details about the algorithm we use, see Appendix D.

[18] Borgatti (2006) calls this the "ensemble problem": The optimal set of nodes for a given optimization problem is not necessarily the set of optimal banks when considered individually.

[19] Even though we are concerned with minimizing welfare losses, we define the net welfare effect such that $\Delta^i w$ is *positive* if bailing out bank i is beneficial.

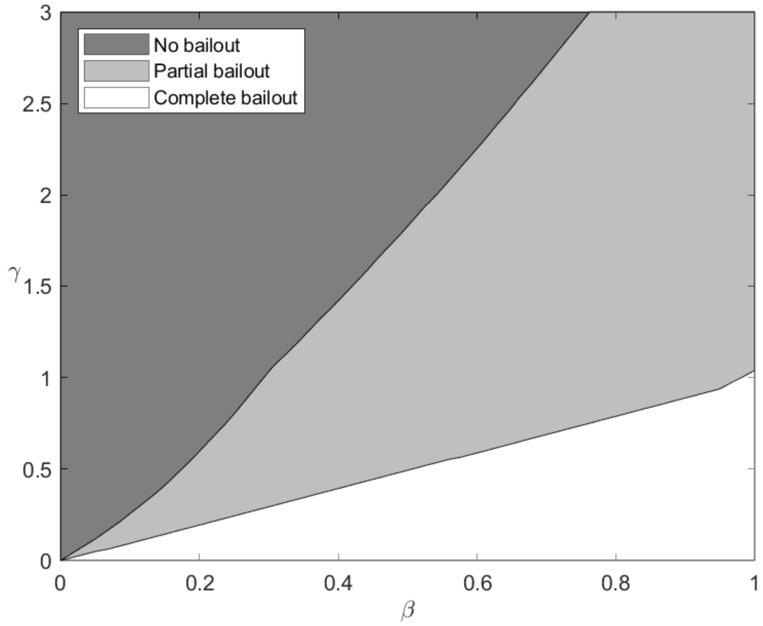


Figure 5: How does the bailout choice depend on cost parameters?

payments from $\bar{p}^i$ to L^i, thus reducing the shortfalls and deadweight losses of i's creditors, their creditors, and so forth. Some troubled banks may even become solvent as a result. On the other hand, the bailout comes at a cost: First, the increase in sovereign debt implies a lower sovereign debt price q_1 and a higher probability of sovereign default; second, the drop in q_1 increases the shortfalls and deadweight losses in other banks that are still insolvent after bailing out bank i. We refer to these two sources of welfare losses, respectively, as γ-component and β-component, named after their respective coefficients in the social welfare function (9). In what follows we first elaborate on all components individually before we state the decomposition in a lemma.

We start with an observation about the benefit of bailing out bank i: For a given set of surviving banks $\mathcal{S}$, by bailing out bank $i \in \mathcal{S}^c$ the government avoids bankruptcy deadweight losses of

$$\beta \sum_{j \in \mathcal{I}} \left[\chi_1^j + (L^j - \bar{p}^j(\mathcal{S})) \left(\sum_k \pi^{kj} + (1+\beta) \sum_k \sum_l \pi^{lk} \pi^{kj} + (1+\beta)^2 \sum_k \sum_l \sum_m \pi^{ml} \pi^{lk} \pi^{kj} + \dots \right) \right], \tag{10}$$

where $\mathcal{I}$ denotes the set of banks that become solvent through the bailout (with $i \in \mathcal{I}$) and all summations except the first are over the default set $\mathcal{S}^c \setminus \mathcal{I}$. The expression in (10) captures (a) the direct impact of avoiding bankruptcy deadweight losses $\beta \chi_1^j$ for banks in $\mathcal{I}$, and (b) the indirect impact of their increased interbank repayments. The idea is that the increase in repayments from $\bar{p}^j(\mathcal{S})$ to L^j not only increases j's creditors' assets available for repayment, but also reduces their shortfall and hence their deadweight losses. For example, if bank j fully repays, its creditor k's shortfall decreases by $\pi^{kj}(L^j - \bar{p}^j(\mathcal{S}))$ and its repayment $\bar{p}^k$ even increases by $(1+\beta)\pi^{kj}(L^j - \bar{p}^j(\mathcal{S}))$, and so on. In other words, the positive effect of the bailout travels through the subnetwork of defaulting banks and is amplified by the bankruptcy cost coefficient β at each node.

The expression in (10) can be further simplified. Observe that the increase in interbank repayments is bounded from above by L^j, namely if bank j originally could not even partially repay its interbank liabilities ($\bar{p}^j(\mathcal{S}) = 0$). Otherwise, that is if $\bar{p}^j(\mathcal{S}) > 0 \ \forall j \in \mathcal{I} : L^j > 0$, we can use Definition 1 to show that $L^j - \bar{p}^j(\mathcal{S}) = (1+\beta)\chi_1^j$. For ease of notation we restrict our subsequent analysis to the latter case, so that the expression in (10) becomes

$$\beta \sum_{j\in\mathcal{I}} \chi_1^j \underbrace{\left(1 + (1+\beta)\sum_k \pi^{kj} + (1+\beta)^2 \sum_k \sum_l \pi^{lk}\pi^{kj} + \dots\right)}_{C^j} \tag{11}$$

We remark, however, that our results would qualitatively also hold in the general case where j's depositors incur losses, i.e., if there exists some bank j with positive interbank liabilities $L^j > 0$ which makes zero payments to its creditors in the network.

Borrowing from Glasserman and Young (2015), we refer to the term in parentheses as the "node depth" of bank j and denote it by C^j which stands for "centrality". Node depth measures the extent to which any loss originating at bank j gets spread and amplified in the subnetwork of defaulting banks. Therefore, it is naturally increasing in the set of defaulting banks, i.e., for a given network Π, node depth of any individual bank is weakly higher if more banks default. Let $\mathcal{D}$ be a set of defaulted nodes, and $\Pi_{\mathcal{D}}$ be the $|\mathcal{D}| \times |\mathcal{D}|$ matrix obtained by restricting the relative liabilities matrix Π to $\mathcal{D}$. Moreover, denote by $I_{\mathcal{D}}$ be the $|\mathcal{D}| \times |\mathcal{D}|$ identity matrix, and let Π' denote the transpose of matrix Π. Conditional on $\mathcal{D}$, the vector of node depths of banks in $\mathcal{D}$ is given by $[I_{\mathcal{D}} - (1+\beta)\Pi'_{\mathcal{D}}]^{-1} \cdot \mathbb{1}_{\mathcal{D}}$ if the spectral radius of $(1+\beta)\Pi'_{\mathcal{D}}$ is less than one. This technical condition has a clear economic interpretation: For node depth to be well defined, the entries of the relative liability matrix of the defaulting banks $\Pi'_{\mathcal{D}}$ have to be sufficiently small, i.e., the defaulting banks need to owe a sufficient fraction of their liabilities to solvent banks outside of $\mathcal{D}$. The higher the bankruptcy cost parameter β, the lower the relative exposures within the defaulting set must be.[20]

Another interpretation is that solvent banks are required to act as a valve to absorb some of the losses in the interbank market. Otherwise there would be too much amplification of welfare losses within $\mathcal{D}$ and the series in (10) and (11) would not converge. For example, if the set of defaulting banks had a *ring* structure with $\pi^{ij} = 1$ if $j = i+1$, modulo n, node

---

[20] We derive formal conditions for two- and three-dimensional default sets in Appendix C.

depth could not be computed using the matrix inversion formula given above. Rather than fading out, any dollar of shortfall at a given bank would be fully transmitted and amplified at every node as it travels through the ring until all assets are wiped out. In contrast, if the interbank network is *complete* and *symmetric* with $\pi^{ij} = \frac{1}{n-1}$ $\forall i, j \in N$, node depth can be computed as long as the total number of banks n exceeds the number of defaulting banks by a sufficient amount. The critical number of banks increases with the bankruptcy cost parameter β.

There is an interesting analogy between the concept of node depth and other, more standard network metrics such as eigenvector or Katz-Bonacich centrality. First, just like eigenvector centrality, a bank's node depth is a weighted sum of the node depths of its creditors, with weights given by the relative liability matrix Π' (recall that π^{kj} measures the share of k's claim in j's total interbank liabilities). Intuitively, a bank's node depth is higher if it has higher liabilities to defaulting banks that have themselves high liabilities to other defaulting banks. And second, similarly to the attenuation coefficient in Katz-Bonacich centrality, the factor $(1 + \beta)$ controls the effect of path length on the relationship between different nodes' centralities. We refer to section 9.3 of Glasserman and Young (2016) for more details.

From an economic perspective, the expression in (11) indicates that the benefit of a bailout is the largest if targeted at banks with (a) large shortfalls, (b) high potential to avoid contagious defaults (i.e., a large set $\mathcal{I}$), and (c) high centrality in the sense of node depth. However, a comprehensive analysis of the government's tradeoff also requires to analyze the cost of a bailout.

We begin by considering the γ-component, i.e., the part of the cost that is directly related to the pecuniary expenditure the bailout of bank i requires. The increase in the level of public debt (and thus in the sovereign default probability) depends on four items: First of all, from equation (8) we can see that, *ceteris paribus*, a larger shortfall χ_1^i clearly requires higher bailout expenditures and thus leads to a sharper increase in the sovereign default probability and a sharper drop in the bond price q_1. Second, there is a doom loop multiplier effect: a higher sovereign exposure b^i also leads to higher bailout expenditures because every dollar raised for the bailout of bank i increases its effective shortfall by an amount proportional to b^i. Therefore, even for identical shortfalls, bailing out banks with more sovereign debt on their balance sheets comes with a higher γ-component than bailing out low-exposure banks. Third, depending on the sovereign exposure of banks in $\mathcal{S}$, it may happen that the drop in q_1 makes additional bailout transfers to other banks necessary to prevent their insolvency. The cost of raising these funds will also be reflected in the γ-component. Fourth and finally, it matters whether or not bank i is part of a cycle in the subnetwork of defaulting banks $\mathcal{S}^c$. If so, then bailing out i eventually also increases the interbank repayments it receives, thereby reducing the effective shortfall the government needs to cover according to (8). Summing up, to minimize the γ-component of the bailout cost, the government would ideally bail out banks that have only a low shortfall to begin with, hold little sovereign debt, and have interbank claims on other defaulting banks that have themselves (direct or indirect) claims on the bailout candidate.

To complete the picture, we turn to the β-component which describes the additional bankruptcy deadweight losses in defaulting banks that are triggered by the drop in the sovereign bond price which we denote by $\Delta^i q_1 < 0$. The bailout of bank i causes deadweight losses in the amount of $|\beta \Delta^i q_1 b^j|$ for each defaulting bank $j \in \mathcal{S}^c \setminus \mathcal{I}$ which will then be amplified in proportion to their centrality C^j. This means that *the side effects of a bailout can become very costly if the remaining defaulters are highly exposed to their government and very central in the subnetwork of defaulting banks.* It is exactly this key observation that will give rise to our main result further below. The following lemma summarizes our discussion of the government's bailout incentives so far.

Lemma 2 (Decomposition of net welfare effect)**.** *For a given set of surviving banks $\mathcal{S}$, the net welfare effect $\Delta^i w$ of bailing out bank $i \in \mathcal{S}^c$ can be written as*

$$\Delta^i w = \underbrace{\beta \sum_{j \in \mathcal{I}} \chi_1^j \times C^j}_{\textit{Benefit}} \underbrace{\underbrace{+\beta \sum_{j \in \mathcal{S}^c \setminus \mathcal{I}} \Delta^i q_1 b^j \times C^j}_{\beta\textit{-component}} \underbrace{-\gamma \Delta^i (B_1 P(Default))}_{\gamma\textit{-component}}}_{\textit{Cost}},$$

where $\Delta^i q_1 < 0$ and $\Delta^i (B_1 P(Default)) > 0$ denote changes implied by the bailout of i.

The lemma allows us to make a number of statements about which banks deliver the highest welfare improvements if bailed out. For example, note that in an economy without bank-held domestic sovereign debt ($b = 0$), the β-component would disappear and banks could be ranked in terms of network spillovers only. More precisely, among two banks i and j with the same shortfall (i.e., $\chi_1^i = \chi_1^j$) and the same set of contagious defaults ($\mathcal{I} \setminus i = \mathcal{J} \setminus j$) the γ-component would be identical for both banks. The government would then prefer to bail out the bank with higher node depth C.[21] Our model thus supports the claim that banks with high liabilities to other weakly capitalized banks can be "too interconnected to fail".

Another thought experiment (again with $b = 0$) helps to emphasize the role of node depth and hence network structure: Consider two banks with the same node depth ($C^i = C^j$), but such that has a larger shortfall without a bailout, e.g. $\chi_1^i > \chi_1^j$. Depending on model parameters (in particular for sufficiently high γ) it is possible that the government prefers to bail out bank j if both banks' centrality is low, but it prefers bank i if their centrality

---

[21] This insight echos a number of results from the more general literature about intervention in networks. For instance, Galeotti *et al.* (2019) show that in network games of strategic complements, the optimal intervention allocated to a given node is proportional to its eigenvector centrality.

is sufficiently high. The reason is that the benefit of a bailout is proportional to the bank's centrality while the fiscal cost (γ-component) is not. Therefore, if centrality is low, the higher γ-component of bailing out i can dominate the government's decision, but with sufficiently high centrality the higher benefit overcompensates the increased sovereign default risk.

In what follows we often compare two banks or two entire financial systems that are identical in all but one dimension. This allows us to isolate specific channels and derive precise analytical results.

Sovereign Debt Distribution and Network Centrality

The insights discussed in the previous paragraphs could, in principle, also be derived in a conventional model of financial contagion without the doom loop. In contrast, this section analyzes how network centrality interacts with the distribution of sovereign debt and hence with the doom loop.

Before we study non-degenerate, dispersed distributions of sovereign debt we consider the simple benchmark case in which every bank has the same exposure to the government.

Proposition 3 (Uniform sovereign debt distribution)**.** *Let* $(L, \Pi, \tilde{c}, d, \tilde{b})$ *and* $(L, \Pi, \hat{c}, d, \hat{b})$ *be two financial systems such that*

1. $\tilde{b} = 0$
2. $\hat{b}^k = b' > 0 \; \forall k \in N$
3. $\hat{c} = \tilde{c} - \hat{b}/R$ *(which implies* $\hat{\chi}_1 = \tilde{\chi}_1 \equiv \chi_1$*).*

Moreover, let $i, j \in \mathcal{S}^c$ *be two banks that differ only in their node depth, with* $C^i > C^j$*. In particular, they have identical shortfalls* $\chi_1^i = \chi_1^j$ *and cause the same set of contagious defaults* $\mathcal{I} \setminus i = \mathcal{J} \setminus j$*. Then we have*

$$\tilde{\Delta}^i w > \tilde{\Delta}^j w, \quad \hat{\Delta}^i w > \hat{\Delta}^j w, \quad \text{and} \quad \hat{\Delta}^i w - \hat{\Delta}^j w > \tilde{\Delta}^i w - \tilde{\Delta}^j w.$$

Proof. See Appendix B. ■

In the proposition, the tilde-system captures the setting described above without any sovereign exposure (condition 1), so there is no β-component to consider. For two banks with identical shortfalls and identical contagious defaults, the difference in node depth $C^i > C^j$ breaks the tie, so the government prefers to bail out bank i. The hat-system, in contrast, features a uniform sovereign debt distribution across banks (conditions 2 and 3). The proposition states that the presence of domestic sovereign exposure on banks' balance sheets - even if it is uniformly distributed - *strenghtens* the role of centrality as a driver of the government's optimal decision. If bailing out bank i is preferred to bailing out j in the scenario without the doom loop, it is also preferred with the doom loop, but not vice versa. The reason is that now, on top of a higher benefit, bailing out bank i is associated with a lower β-component than bailing out bank j (remember that the β-component is increasing in defaulting banks' centrality). Put differently, even a highly symmetric sovereign debt distribution amplifies the "too interconnected to fail" problem.

Equipped with this general result on the interaction of centrality and sovereign exposure, we next consider asymmetric distributions of domestic sovereign debt. *Does the government prefer to bail out banks with more or less domestic sovereign debt?* From Lemma 2 we know that domestic sovereign bonds on a bank's balance sheet have two counteracting effects: On the one hand, high sovereign exposure b^i makes it financially costly to bail out bank i because every bailout dollar transferred to bank i lowers q_1 and hence further increases i's shortfall. This "doom loop multiplier" effect (which translates into a high γ-component) is stronger for larger b^i. On the other hand, high sovereign exposure also generates welfare losses through the β-component if bank i is not bailed out, especially if i is very central in terms of node depth. In other words, a higher b^i increases the effective shortfall of bank i and hence the required bailout expenditures, but it can also make it extremely socially costly *not* to save bank i in a partial bailout.

Therefore, whether a given bank $i \in \mathcal{B}$ will optimally be bailed out depends on the relative strength of these two forces. Crucially, whereas the first effect applies to each bank equally, the second effect is proportional to a bank's network centrality. In particular, a bank with high centrality and sovereign exposure may be more likely to be bailed out than an identical bank with little or no exposure. In contrast, peripheral banks (for which the second effect is negligible) are unambiguously less likely to be saved the more sovereign bonds they hold. We formalize this main result in the following proposition.

Proposition 4 (Sovereign debt distribution within networks). *Let $i, j \in \mathcal{S}^c$ be two banks that differ only in their asset composition ($b^i > b^j$ and $c^i < c^j$). In particular, they have identical shortfalls $\chi_1^i = \chi_1^j \equiv \chi_1$ and centrality $C^i = C^j \equiv C$ and cause the same set of contagious defaults $\mathcal{I} \setminus i = \mathcal{J} \setminus j$. Then $\exists C^*$ such that*

$$\Delta^i w > \Delta^j w \quad \textit{iff} \quad C \geq C^*$$

as long as $\frac{b^i}{b^j} > \frac{\Delta^i q_1}{\Delta^j q_1} > 1$. The threshold C^* is increasing in γ and decreasing in β.

Proof. See Appendix B. ■

If the government has to choose between two otherwise identical banks, whether it will pick the one with more or less domestic sovereign exposure depends on the banks' centrality

Table 2: Balance sheet structure of banks in the network

	1	2	3	4	5	6	7
c^i	8.22	10.8	19.4	22.0	22.0	31.0	31.0
b^i	5.33	2.67	5.33	2.67	2.67	2.67	2.67
$(\pi L)^i$	20.2	20.2	8.96	8.96	8.96	0	0
Total assets	33.6	33.6	33.6	33.6	33.6	33.6	33.6
d^i	30.3	30.3	16.8	16.8	16.8	16.8	16.8
L^i	0	0	13.4	13.4	13.4	13.4	13.4
Total liabilities	30.3	30.3	30.3	30.3	30.3	30.3	30.3
Equity	3.36	3.36	3.36	3.36	3.36	3.36	3.36

Note: Banks 1 and 3 hold more sovereign debt than their counterparts.

C. The reason is that the welfare losses due to the β-component (which are higher if bank j with the smaller sovereign exposure is saved) are proportional to centrality C, as shown in Lemma 2. Hence, for large C the difference in β-components dominates the difference in the γ-components so that bailing out bank i becomes more attractive. In contrast, for peripheral banks (low C) the β-component carries only little weight in the welfare comparison and the lower γ-component leads the government to bail out bank j.

We illustrate the result through a stylized numerical example with $n = 7$ banks: Suppose the government initially owes a debt stock of $B_0 = 120$ (which may be interpreted as 120% of GDP), of which 20% is held by the domestic banking sector, so $\sum_{i \in N} b^i = 24$. The gross risk-free interest rate is $R = 1.02$, hence the initial sovereign bond price is given by $q_0 = R^{-1} = 0.9804$. Moreover, we pick the shape parameter of the (Pareto) tax capacity distribution to be $\alpha = 0.9$ which delivers quantitatively reasonable movements in sovereign spreads.

Table 2 summarizes banks' initial balance sheets (i.e. before the unexpected fundamental shock happens). Total balance sheet size is chosen such that the average bank holds 10% of its balance sheet in domestic sovereign debt, just like Italy in Figure 1. To maintain transparency and simplicity, we initially let every bank have the same balance sheet size and leverage (equity buffer = 10% of total assets). Note, however, that banks 1 and 3 each hold twice as much sovereign debt as the remaining banks.

The interbank network is visualized in Figure 6. Banks 1 and 2 have no interbank liabilities, but are lending to three intermediaries (3,4,5) who further lend to banks 6 and 7. The squares mark the set of fundamentally defaulting banks $\mathcal{F} = \{1, 2, 3, 5\}$ whose "cash" assets are reduced by the shock such that their total assets drop by 20%, leaving them with a shortfall of 10% each.[22] There is no default contagion because none of the fundamentally defaulting banks have any liabilities to banks outside of $\mathcal{F}$. Summing up, this is the simplest possible setup to demonstrate the content of the proposition; we have two sets of almost identical failing banks (1 vs. 2 and 3 vs. 5) that only differ in terms of sovereign exposure

---

[22] To make the local analysis interesting, we picked the deadweight loss parameters $\beta = 0.8$ and $\gamma = 0.7$ such that the optimal bailout is a *partial* bailout of banks 3 and 5, i.e. banks 1 and 2 default in equilibrium.

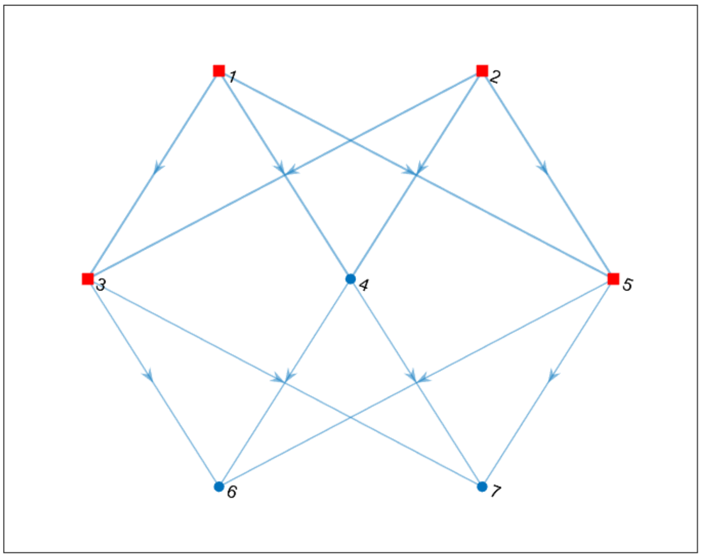


Figure 6: Banks 3-5 borrow from 1 and 2 and lend to 6 and 7 (squares represent $\mathcal{F}$)

within pairs. In particular, banks 1 and 2 have a trivial node depth of $C^1 = C^2 = 1$ (because they do not have any interbank liabilities) whereas banks 3 and 5 have $C^3 = C^5 = 2.8$. Just like the proposition states, among the peripheral banks the government prefers bailing out the low-b bank 2 ($\Delta^2 w = -0.82$) to the high-b bank 1 ($\Delta^1 w = -0.86$), but among the more central banks it prefers the high-b bank 3 ($\Delta^3 w = 4.89$) to the low-b bank 5 ($\Delta^5 w = 4.79$).

Finally, note that our proposition requires that the difference in sovereign exposures must overcompensate the difference in absolute price changes, i.e. b^i and b^j must be sufficiently different. Otherwise the fact that saving bank i has a stronger price impact ($|\Delta q_1^i| > |\Delta q_1^j|$) dominates the β-component so that the government unambiguously prefers to save bank j, regardless of centrality. This caveat suggests that our result cannot be interpreted as the *marginal* effect of higher sovereign exposure on the bailout outcome.

One interpretation of the result in Proposition 4 is that the doom loop alters the government's ranking of banks to be rescued. Whereas in a world without the doom loop (e.g. with $b = 0$), the government would be indifferent between two banks with the same shortfall χ_1 and node depth C, the different exposures to the doom loop now break the tie. In other words, network centrality alone is not sufficient anymore to pin down the optimal set of bailed-out banks. This highlights the doom loop as a new channel of contagion that conventional centrality measures cannot capture.

Another way of looking at the effect of holding more domestic sovereign debt is *across* networks (whereas Proposition 4 is based on a *within*-network comparison). To this end we consider two financial systems that are identical in every regard, except that an arbitrary defaulting bank *i* holds more sovereign debt in one system that in the other (conditions 1 and 2). To keep the overall surplus of the system unchanged, suppose that this bank holds less cash in return so that its overall shortfall remains unchanged (condition 3). The following proposition shows that whether the higher b^i in the second system increases or decreases bank *i*'s position in the government's "bailout ranking" depends on its centrality.

Proposition 5 (Sovereign debt distribution across networks). *For a fixed B_0, let $(L, \Pi, \tilde{c}, d, \tilde{b})$ and $(L, \Pi, \hat{c}, d, \hat{b})$ be two financial systems such that*

1. $\exists ! i \in \mathcal{S}^c : \hat{b}^i > \tilde{b}^i$
2. $\hat{b}^j = \tilde{b}^j$ *and* $\hat{c}^j = \tilde{c}^j \; \forall j \neq i$
3. $\hat{c}^i = \tilde{c}^i - (\hat{b}^i - \tilde{b}^i)/R.$

Then $\forall j \neq i \; \exists C^*(j)$ such that

$$\hat{\Delta}^i w - \hat{\Delta}^j w > \tilde{\Delta}^i w - \tilde{\Delta}^j w \quad \textit{iff} \quad C^i \geq C^*(j).$$

The threshold is increasing in γ and decreasing in β, $|\hat{\Delta}^j q_1|$ and $(\hat{b}^i - \tilde{b}^i)$.

Proof. See Appendix B. ■

The above proposition can be easily explained thanks to the decomposition in Lemma 2. In the hat-network bailing out bank *i* results in a larger β-component and a larger γ-component than in the tilde-network because of the doom loop multiplier and the corresponding larger price impact. Meanwhile, the benefit component is the same in both networks. As a consequence, the absolute net benefit of bailing out *i* is clearly lower in the hat-network, i.e. $\hat{\Delta}^i w < \tilde{\Delta}^i w$. However, all other bailout candidates *j* see their β-component increase (and thus their $\Delta^j w$ fall) because the defaulting *i* now holds more sovereign debt. Moreover, this increase in the β-component is an increasing function of *i*'s centrality. Hence, if *i* is central enough it can even end up higher in the government's ranking than before.

Summing up, our results show that a higher b^i does not make a bailout of *i* more attractive in absolute terms, nor does it increase the government's propensity to engage in a bailout in the first place. However, it can make bailing out other banks even less attractive, namely if *i* is sufficiently central in the interbank network.

3.5 Sovereign Debt Exposure: A Strategic Tool?

Even though we do not model banks' optimal behavior explicitly, the results outlined in the previous section have important implications for banks' strategic portfolio choice *ex ante*. We have shown that, depending on a bank's centrality, higher sovereign exposure can make a bailout of that bank more likely *ex post*. Hence, if banks value the prospect of a bailout and anticipate the government's best response, they can use their sovereign debt portfolio to affect the odds of being bailed out.

One way to illustrate this in the context of Proposition 4 is the following: Suppose that banks wake up in $t = -1$, knowing their position in the interbank network (and hence their node depth for each shock realization), but having not yet chosen their domestic sovereign exposure. In terms of Proposition 4, a given bank can choose to be either bank i or bank j. Then, banks with high centrality can increase the chance to be bailed out by purchasing more sovereign debt than otherwise identical banks. Moreover, the fact that the threshold C^* is decreasing in β implies that this result applies to more banks in financial systems where bankruptcy costs are higher.

Put differently, our results suggest that systemically important banks may have an incentive to load up on domestic sovereign debt to drive up the cost of *not* bailing them out when the chips are down.[23]

4 Conclusion

In this paper we study the government's optimal bailout strategy in a banking crisis if (a) financing a bailout depresses the value of domestic sovereign debt on bank balance sheets (the "doom loop") and (b) banks are connected to each other through a network of liabilities.

We construct a stylized banking network model in which banks hold domestic sovereign debt on their balance sheets. An adverse shock renders a subset of banks insolvent. Without government intervention, these bank failures can spread to (direct or indirect) creditor banks and cause deadweight losses in the banking sector. To reduce these deadweight losses, the government can borrow funds in the sovereign debt market and distribute them optimally to failing banks. However, an increase in the stock of public debt makes it more risky and leads to a higher discount on sovereign bonds. This weakens banks' balance sheets further, adding to the cost of the bailout.

We find, not surprisingly, that the doom loop is weaker (and bailouts therefore less costly) if banks hold less domestic sovereign debt or if banks with large domestic sovereign exposure are well capitalized. This result is directly related to the current debate about higher regulatory capital charges for sovereign exposure in the eurozone (see Véron (2017)).

---

[23] Of course, this is just one possible explanation. In fact, there are many more factors driving banks' sovereign home bias, from rather obvious ones like capital and liquidity regulation to more subtle motives such as the "risk-shifting" argument in Crosignani (2017). Which explanation is relevant in any particular country is ultimately an empirical question. Unfortunately, due to their highly confidential nature, robust data on bilateral interbank exposures are notoriously hard to obtain.

Moreover, we show that the government can rank otherwise identical banks according to their "node depth", a centrality measure introduced by Glasserman and Young (2015). Compared to a situation without sovereign debt on banks' balance sheets, even a uniform distribution of sovereign debt across banks strengthens the role of centrality as a tiebreaker and thus exacerbates the "too-interconnected to fail" problem.

Our main result is that the optimal subset of banks to bail out depends jointly on their position in the interbank network and their domestic sovereign exposure. In particular, if banks with high amounts of sovereign bonds are systemically important (e.g. because they have lots of fragile creditors), the government may prefer to save them rather than banks with low domestic sovereign exposure, even though this requires a larger bailout. Equivalently, if there is a bailout, highly central banks are more likely to be part of it if they hold more sovereign debt since that increases the cost of letting them fail. While higher sovereign exposure unambiguously makes a bailout of a given bank more expensive, it can make bailing out other banks instead even more costly.

Extrapolating from this finding, we argue that a bank's position in the network may have strategic implications for its individually optimal domestic sovereign exposure. Systemically important banks might be able to use domestic sovereign debt as a "strategic tool" to increase the likelihood of being bailed out. Our model therefore provides a new, network-based perspective on the question of why banks in stressed European countries increased their domestic sovereign bond holdings during the European sovereign debt crisis.

Our results contribute to the understanding of (a) the government's tradeoff in the face of a banking crisis and (b) banks' incentives to exploit the government's willingness to bail them out. On a more general level, our study is part of a broader research agenda that shows where conventional models with representative banks understate or even miss important effects altogether. Moreover, our model can be used as an augmented stress testing framework with endogenous sovereign risk that feeds back to the banking system. Hence, it can support the policy making process of regulatory and supervisory agencies such as the European Banking Authority (EBA) or the ECB.

Finally, our model results point to an interesting political economy interpretation. On the one hand, to the extent that governments rely on domestic banks as a cheap and reliable source of financing, policymakers may want to refrain from limiting their systemically important banks' domestic sovereign bond holdings. On the other hand, if a banking crisis were to occur in the future, these systemically important banks are precisely the institutions that would require expensive bailouts. As a consequence, there is a tension between the incentives of current and future governments. Our paper only considers the second effect and hence the regulatory objective of minimizing future risks. In other words, our model only captures the long-term benefits of limiting domestic sovereign exposure, but not the potential short-term cost.

Our framework may be extended along several interesting dimensions. In one possible variation, bank failures would prompt fire sales of sovereign debt to risk-averse investors that lead to an increase in sovereign spreads. As a consequence, the government would now have a clear incentive to rescue the banks with large stocks of domestic sovereign debt, regardless of their network centrality.

A second, more ambitious extension is to develop a fully fledged game theoretical model, in which banks strategically choose their sovereign bond holdings by maximizing an exogenously specified objective function. Banks would thus endogenously construct their balance sheet in anticipation of government intervention to maximize their bailout option. In the present paper, we focus on *ex post* optimal bailouts and do not allow for strategic interactions, so our model ignores the role of fiscal commitment. If the government in our model was committed to a strict no-bailout policy, the doom loop (and hence the "strategic tool" interpretation of bank purchases of sovereign bonds) would disappear. In a richer model, a credible no-bailout policy would affect banks' risk-taking and sovereign debt choice *ex ante* as in Cooper and Nikolov (2018) and could lead to an overall improvement in welfare. However, it is hard to conceive of a credible no-bailout commitment that would remain intact even under the most catastrophic circumstances.[24]

Third and finally, our model deliberately abstracts from the "real economy loop" in Brunnermeier *et al.* (2016) which is already well understood. That means we do not consider the fact that bank failures could also affect the sovereign's creditworthiness *without* bailouts through a reduction in lending, an investment slump, and reduced tax capacity. Incorporating this channel in our model would increase the direct macroeconomic social benefits of a bailout *ceteris paribus*. A similar channel would work in the opposite direction and make bailouts less attractive: Rising sovereign spreads (e.g. triggered by a bailout) tend to be passed through to corporate spreads and household borrowing rates. This "crowding out" effect, documented for example in Demirci *et al.* (2019), might eventually lead to lower tax revenue and thus even higher spreads.

In general, whether for any particular economy our model over- or understates the effects of the doom loop vis-à-vis representative bank models such as Brunnermeier *et al.* (2016) depends on the empirical size of interbank lending relative to domestic sovereign bond portfolios. If interbank lending is relatively insignificant, our model overstates the doom loop effect and representative bank models suffice. In the opposite case, however, our paper adds an important perspective that goes unnoticed in one-bank models.

---

[24] Note, for example, that in the face of the Covid-19 pandemic many governments (including in the US and EU) have rolled back no-bailout commitments that were enacted in the aftermath of the 2008 crisis.

References

ABAD, J. (2019). Breaking the Sovereign-Bank Nexus. Mimeo.

ACEMOGLU, D., OZDAGLAR, A. and TAHBAZ-SALEHI, A. (2015). Systemic Risk and Stability in Financial Networks. *The American Economic Review*, **105** (2), 564–608.

ACHARYA, V., DRECHSLER, I. and SCHNABL, P. (2014). A Pyrrhic Victory? Bank Bailouts and Sovereign Credit Risk. *The Journal of Finance*, **69** (6), 2689–2739.

ACHARYA, V. V. and STEFFEN, S. (2015). The "Greatest" Carry Trade Ever? Understanding Eurozone Bank Risks. *Journal of Financial Economics*, **115** (2), 215–236.

ADRIAN, T. and SHIN, H. (2010). Liquidity and Leverage. *Journal of Financial Intermediation*, **19** (3), 418–437.

ALLEN, F. and GALE, D. (2000). Financial Contagion. *Journal of Political Economy*, **108** (1), 1–33.

ALTAVILLA, C., PAGANO, M. and SIMONELLI, S. (2017). Bank Exposures and Sovereign Stress Transmission. *Review of Finance*, **21** (6), 2103–2139.

ALTINOGLU, L. and STIGLITZ, J. E. (2019). Collective Moral Hazard and the Interbank Market. Mimeo.

BERNARD, B., CAPPONI, A. and STIGLITZ, J. E. (2018). Bail-Ins and Bail-Outs: Incentives, Connectivity, and Systemic Stability. *NBER Working Paper No. 23747*.

BERNSTEIN, S., COLONNELLI, E. and IVERSON, B. C. (2019). Asset Allocation in Bankruptcy. *The Journal of Finance*, **74** (1), 5–53.

BORGATTI, S. P. (2006). Identifying Sets of Key Players in a Social Network. *Computational & Mathematical Organization Theory*, **12** (1), 21–34.

BOZ, E., D'ERASMO, P. and DURDU, C. B. (2014). Sovereign Risk and Bank Balance Sheets: The Role of Macroprudential Policies. *Manuscript, International Monetary Fund*.

BRUNNERMEIER, M. K., GARICANO, L., LANE, P. R., PAGANO, M., REIS, R., SANTOS, T., THESMAR, D., VAN NIEUWERBURGH, S. and VAYANOS, D. (2016). Breaking the Sovereign-Bank Diabolic Loop: A Case for ESBies. In *American Economic Review Papers and Proceedings*.

CABRALES, A., GALE, D. and GOTTARDI, P. (2016). Financial Contagion in Networks. In *The Oxford Handbook of the Economics of Networks*.

—, GOTTARDI, P. and VEGA-REDONDO, F. (2017). Risk Sharing and Contagion in Networks. *The Review of Financial Studies*, **30** (9), 3086–3127.

CALVO, G. A. (1988). Servicing the Public Debt: The Role of Expectations. *The American Economic Review*, pp. 647–661.

COLE, H. L. and KEHOE, T. J. (2000). Self-fulfilling Debt Crises. *The Review of Economic Studies*, **67** (1), 91–116.

COOPER, R. and NIKOLOV, K. (2018). Government Debt and Banking Fragility: The Spreading of Strategic Uncertainty. *International Economic Review*, **59** (4), 1905–1925.

CROSIGNANI, M. (2017). Why are Banks not Recapitalized During Crises? FEDS Working Paper 2017-084.

DEMIRCI, I., HUANG, J. and SIALM, C. (2019). Government Debt and Corporate Leverage: International Evidence. *Journal of Financial Economics*, **133** (2), 337–356.

EISENBERG, L. and NOE, T. H. (2001). Systemic Risk in Financial Systems. *Management Science*, **47** (2), 236–249.

ELLIOTT, M., GOLUB, B. and JACKSON, M. O. (2014). Financial Networks and Contagion. *The American Economic Review*, **104** (10), 3115–3153.

EROL, S. (2019). Network Hazard and Bailouts. Mimeo.

FARHI, E. and TIROLE, J. (2018). Deadly Embrace: Sovereign and Financial Balance Sheets Doom Loops. *The Review of Economic Studies*, **85** (3), 1781–1823.

GABALLO, G. and ZETLIN-JONES, A. (2016). Bailouts, Moral Hazard and Bank's Home Bias for Sovereign Debt. *Journal of Monetary Economics*, **81**, 70–85.

GALEOTTI, A., GOLUB, B. and GOYAL, S. (2019). Targeting Interventions in Networks. Mimeo.

GLASSERMAN, P. and YOUNG, H. P. (2015). How Likely is Contagion in Financial Networks? *Journal of Banking & Finance*, **50**, 383–399.

— and — (2016). Contagion in Financial Networks. *Journal of Economic Literature*, **54** (3), 779–831.

GREENWALD, B. C. and STIGLITZ, J. E. (1993). Financial Market Imperfections and Business Cycles. *The Quarterly Journal of Economics*, **108** (1), 77–114.

GREENWOOD, R., LANDIER, A. and THESMAR, D. (2015). Vulnerable Banks. *Journal of Financial Economics*, **115** (3), 471–485.

KIYOTAKI, N. and MOORE, J. (1997). Credit Chains. Mimeo.

POPOV, A. and VAN HOREN, N. (2015). Exporting Sovereign Stress: Evidence from Syndicated Bank Lending during the Euro Area Sovereign Debt Crisis. *Review of Finance*, **19** (5), 1825–1866.

THALER, D. (2018). Sovereign Default, Domestic Banks and Exclusion from International Capital Markets. Banco de España Working Paper.

VÉRON, N. (2017). Sovereign Concentration Charges: A New Regime for Banks' Sovereign Exposures. In *Study for the European Parliament, Brussels. IDEAS FROM THE 'FUTURE EUROPE' FORUM*, vol. 67.

Appendix

A Assumptions about the Tax Capacity Function

We want the graph of the function on the right-hand side of equation (4) to (a) start above the 45° line, and (b) to be strictly concave in q_1. Since the price is bounded in $[0, R^{-1}]$, these two conditions guarantee a single intersection with the 45° line and hence a unique fixed point of equation (4). In the following, we derive general conditions on the tax capacity function $F(\cdot)$ that deliver these two results and we show that the Pareto distribution with shape parameter $\alpha < 1$ satisfies these conditions.

Property (a) requires that

$$\lim_{q_1 \to 0} \frac{\partial \frac{1-F\left(B_0 + \frac{T(q_1)}{q_1}\right)}{R}}{\partial q_1} > 1 \tag{12}$$

Taking the derivative with respect to q_1 (ignoring the fact that $T(q_1)$ is piecewise linear and therefore not differentiable everywhere) yields

$$\lim_{q_1 \to 0} -\frac{1}{R}\left(T'(q_1)q_1 - T(q_1)\right)\frac{f\left(B_0 + \frac{T(q_1)}{q_1}\right)}{q_1^2} > 1,$$

where $f(\cdot)$ is the density associated with cdf $F(\cdot)$. This can be rewritten as

$$-\frac{1}{R}\underbrace{\lim_{q_1\to 0}\left(T'(q_1)q_1 - T(q_1)\right)}_{=-T(0)} \times \lim_{q_1\to 0}\frac{f\left(B_0+\frac{T(q_1)}{q_1}\right)}{q_1^2} > 1,$$

where we are using the fact that the limit of a product is equal to the product of the limits if one limit is non-zero and the other infinite or if both are finite. Since the first term converges to a positive constant for $q_1 \to 0$, it is sufficient for the statement to be true if the second limit is infinite, i.e. if

$$\lim_{q_1\to 0}\frac{f\left(B_0+\frac{T(q_1)}{q_1}\right)}{q_1^2} = \infty \iff \lim_{x\to\infty} f(x)x^2 = \infty \tag{13}$$

In other words, for the tax capacity distribution to satisfy (12) it needs to follow a *power law* with sufficiently slow polynomial decay.

As an example, consider the standard Pareto function with the pdf $f(x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}$ where x_m is the minimum value that x can take (B_0 in this paper) and $\alpha > 0$ the so-called shape parameter. The above requirement now reads

$$\lim_{x\to\infty}\alpha B_0^\alpha \frac{x^2}{x^{\alpha+1}} = \infty$$

Note that for $\alpha < 1$, our sufficient condition in (13) holds whereas for $\alpha > 1$ the limit is zero which would violate (12). For $\alpha = 1$ the limit is finite at B_0 which also contradicts (13), but $\frac{T(0)}{R}B_0 > 1$ is still possible, so (12) might still be satisfied. Hence, $\alpha < 1$ is sufficient, but not necessary for (12) to hold.

For property (b), strict concavity, we need

$$\frac{\partial \frac{1-F\left(B_0+\frac{T(q_1)}{q_1}\right)}{R}}{\partial q_1} > 0 \quad \text{and} \quad \frac{\partial^2 \frac{1-F\left(B_0+\frac{T(q_1)}{q_1}\right)}{R}}{\partial q_1^2} < 0. \tag{14}$$

The first expression is equal to

$$\underbrace{-\frac{1}{R}\left(T'(q_1)q_1 - T(q_1)\right)}_{>0}\underbrace{\frac{f\left(B_0+\frac{T(q_1)}{q_1}\right)}{q_1^2}}_{>0} > 0$$

because $f(\cdot)$ is a probability density function (non-negative) and $T(q_1)$ is non-negative and decreasing in q_1. In the following, we derive a sufficient condition on $f(\cdot)$ such that the second derivative in (14) is negative. Differentiating the previous expression again with respect to q_1 yields the requirement that $\forall q_1 \in [0, R^{-1}]$

$$-\frac{1}{R}\left\{\frac{T''(q_1)q_1^3 - 2q_1\left(T'(q_1)q_1 - T(q_1)\right)}{q_1^4} f\left(B_0 + \frac{T(q_1)}{q_1}\right) \right.$$
$$\left. + \left(\frac{T'(q_1)q_1 - T(q_1)}{q_1^2}\right)^2 f'\left(B_0 + \frac{T(q_1)}{q_1}\right)\right\} < 0$$

Since from (2) $T(q_1)$ is a piecewise linear function of q_1, its second derivative is zero (we ignore again the non-differentiable points), so the previous expression can be simplified to

$$-2\left(T'(q_1)q_1 - T(q_1)\right) f\left(B_0 + \frac{T(q_1)}{q_1}\right) + \frac{\left(T'(q_1)q_1 - T(q_1)\right)^2}{q_1} f'\left(B_0 + \frac{T(q_1)}{q_1}\right) > 0$$

Dividing by the term in brackets yields

$$-2f\left(B_0 + \frac{T(q_1)}{q_1}\right) + \frac{T'(q_1)q_1 - T(q_1)}{q_1} f'\left(B_0 + \frac{T(q_1)}{q_1}\right) < 0,$$

where the inequality sign has changed because the divisor is negative. This is equivalent to

$$\frac{2}{q_1} > \frac{T'(q_1)q_1 - T(q_1)}{q_1^2} \frac{f'\left(B_0 + \frac{T(q_1)}{q_1}\right)}{f\left(B_0 + \frac{T(q_1)}{q_1}\right)} = \frac{\partial}{\partial q_1} log\, f\left(B_0 + \frac{T(q_1)}{q_1}\right),$$

which has to hold for all $q_1 \in [0, R^{-1}]$. Rearranging finally yields the necessary condition

$$-\frac{f'\left(B_0 + \frac{T(q_1)}{q_1}\right)}{f\left(B_0 + \frac{T(q_1)}{q_1}\right)} < \psi(q_1) \equiv \frac{2q_1}{\underbrace{T(q_1) - T'(q_1)q_1}_{>0}} \quad \forall q_1 \in [0, R^{-1}]$$

For q_1 such that $f'(\cdot) > 0$ the condition clearly holds true because $f(\cdot) > 0$. However, for q_1 such that $f'(\cdot) < 0$ (i.e., in particular for low q_1 because $\lim_{x\to\infty} f(x) = 0$) we need the fraction on the left-hand side to be small enough. In other words, the decay of the density needs to be sufficiently slow whenever the density decreases (not just asymptotically).

Again, let us verify condition (14) for the Pareto distribution. Recall that the pdf of a Pareto distributed random variable x takes the form $f(x) = \frac{\alpha B_0^\alpha}{x^{\alpha+1}}$ in our case. We have to make sure that

$$\frac{\partial}{\partial q_1} log\, f\left(B_0 + \frac{T(q_1)}{q_1}\right) < \frac{2}{q_1} \quad \forall q_1 \in [0, R^{-1}].$$

For the Pareto distribution this condition becomes

$$\frac{\partial}{\partial q_1} log \left[\alpha B_0^{\alpha} \left(B_0 + \frac{T(q_1)}{q_1} \right)^{-(\alpha+1)} \right] < \frac{2}{q_1} \quad \forall q_1 \in [0, R^{-1}]$$

which can be simplified to

$$-(\alpha+1) \frac{\partial}{\partial q_1} log \left(B_0 + \frac{T(q_1)}{q_1} \right) < \frac{2}{q_1} \quad \forall q_1 \in [0, R^{-1}]$$

Taking the derivative and simplifying yields

$$\frac{T(q_1) - T'(q_1) q_1}{B_0 q_1 + T(q_1)} = \frac{\frac{T(q_1)}{q_1} - T'(q_1)}{B_0 + \frac{T(q_1)}{q_1}} < \frac{2}{\alpha+1} \quad \forall q_1 \in [0, R^{-1}]$$

which, using (3), can be rewritten as

$$\frac{(B_1 - B_0) - T'(q_1)}{B_1} < \frac{2}{\alpha+1} \quad \forall q_1 \in [0, R^{-1}]$$

Finally, we simplify and get

$$\underbrace{B_0 + T'(q_1)}_{>0} > \frac{\alpha-1}{\alpha+1} B_1 \quad \forall q_1 \in [0, R^{-1}]$$

because the change in aggregate bailouts is equal to $T'(q_1) = -\sum_{i \in \mathcal{D}(\bar{p})} b^i$ which is strictly smaller in absolute value than the overall initial level of outstanding sovereign debt B_0 by assumption. Therefore, $\alpha < 1$ is a sufficient condition for the Pareto distribution to satisfy (14).

B Proofs

Proof of Comparative Static Results in Table 1

To show how the bailout space Ψ depends on a specific model parameter θ *ceteris paribus* (for a constant initial shortfall in the banking sector) we simply sign the partial derivatives $\frac{\partial w^B}{\partial \theta}$ and $\frac{\partial w(\bar{p})}{\partial \theta}$. By definition of Ψ, whenever $\frac{\partial w^B}{\partial \theta} > \frac{\partial w(\bar{p})}{\partial \theta}$ we have that the bailout space depends negatively on θ and vice versa.

- $\theta = \alpha$:

  First note that the shape of the tax capacity distribution does not affect $w(\bar{p})$, so how α

affects Ψ depends entirely on its effect on w^B. Suppose $1 > \alpha' > \alpha$, so from Assumption 1 we see that $F_{\alpha'}(\tau) \geq F_\alpha(\tau)\ \forall \tau$ with strict inequality for $\tau > B_0$. In other words, F_α first-order stochastically dominates $F_{\alpha'}$. Then we have

$$\frac{1 - F_{\alpha'}\left(B_0 + \frac{T(q_1)}{q_1}\right)}{R} < \frac{1 - F_\alpha\left(B_0 + \frac{T(q_1)}{q_1}\right)}{R} \quad \forall q_1.$$

Assumption 1 ensures that the fixed point solution of equation (4) satisfies $q_1^{B,\alpha} > q_1^{B,\alpha'}$ in equilibrium, and equivalently $P(Default)_\alpha < P(Default)_{\alpha'}$. Finally, this implies from equation (7) that $w_\alpha^B < w_{\alpha'}^B$. It follows that an increase in α decreases the bailout space Ψ. ■

- $\theta = \gamma$:
  First, note from equation (6) that $\frac{\partial w(\bar{p})}{\partial \gamma} = 0$, so how γ affects Ψ depends entirely on $\frac{\partial w^B}{\partial \gamma} = \left(B_0 + \frac{T(q_1^B)}{q_1^B}\right) F\left(B_0 + \frac{T(q_1^B)}{q_1^B}\right) > 0$. Hence, $\frac{\partial w^B}{\partial \gamma} > 0 = \frac{\partial w(p)}{\partial \gamma}$. ■

- $\theta = \beta$:
  First, note from equation (7) that $\frac{\partial w^B}{\partial \beta} = 0$. Hence, the dependence of Ψ on β is through $\frac{\partial w(\bar{p})}{\partial \beta}$ only. From Definition 1 we know that a higher β causes larger deadweight losses in each defaulting bank, is associated with lower interbank payments $\bar{p}^i$ and therefore with a weakly larger set of defaulting banks $\mathcal{D}(\bar{p})$. Hence the expression in equation (6) is increasing in β, so $\frac{\partial w(\bar{p})}{\partial \beta} > 0$. ■

- $\theta = R$:
  Note that as we R increases, the right-hand side of the pricing equation (4) decreases for every q_1. In particular, this is also true for the equilibrium price (as the fixed point solution of (4)), so we have $\frac{\partial q_1^B}{\partial R} < 0$. Since $\frac{\partial T(q_1^B)}{\partial q_1^B} < 0$ and $F'(\cdot) > 0$, we know from equation (7) that $\frac{\partial w^B}{\partial R} > 0$.

  At the same time, however, note that $\frac{\partial q_0}{\partial R} = \frac{\partial R^{-1}}{\partial R} = -R^{-2} < 0$. Because $\frac{\partial \chi_0^i}{\partial q_0} \leq 0\ \forall i \in N$ with strict inequality $\forall i \in \mathcal{F}$ it follows that $\frac{\partial \chi_0^i}{\partial R} > 0\ \forall i \in \mathcal{F}$ and hence $\frac{\partial w(\bar{p})}{\partial R} > 0$. Whether Ψ increases or decreases with R is therefore ambiguous. ■

Proof of Lemma 1

Note that the equilibrium sovereign debt price q_1 and the level of aggregate bailouts $T(q_1)$ are simultaneously determined by (4) and (2) with $T(q_1) = \sum t^i(q_1)$. Hence, combining these equations yields

$$Rq_1 = 1 - F\left(B_0 + \frac{x - zq_1}{q_1}\right) \tag{15}$$

as the defining equilibrium condition, where $z \equiv \sum_i b^i$ and $x \equiv \sum_i L^i + d^i - c^+ - (\pi L)^i$ and both sums are over the set of banks that require positive bailouts according to (2). Using the Pareto distribution for $F(\cdot)$, the equation becomes

$$Rq_1 = \left(\frac{B_0}{B_0 - z + \frac{x}{q_1}} \right)^{\alpha} \tag{16}$$

In the following, we will show that the equilibrium price q_1 must be decreasing in x, the "autonomous" component of aggregate bailouts.

From our assumptions about $F(\cdot)$ (in particular the Pareto assumption with $\alpha < 1$) we know that for any $x \geq 0$ there is a unique $q_1 \in (0, R^{-1})$ solving the above equation. Note also that the monotonicity of $x \mapsto q_1$ is the same as that of $q_1 \mapsto x$. But the latter mapping can be solved explicitly: From (16)

$$x = \frac{B_0}{R^{\frac{1}{\alpha}}} q_1^{-\frac{1-\alpha}{\alpha}} - (B_0 - z) q_1$$

The mapping $q_1 \mapsto x$ is continuous and decreasing in the interval $(0, R^{-1})$. The continuity follows immediately from the fact that the functions $q_1^{-\frac{1-\alpha}{\alpha}}$ and q_1 are continuous. The fact that it is decreasing follows from the fact that $q_1^{-\frac{1-\alpha}{\alpha}}$ is decreasing in q_1 because $\alpha < 1$, and the function $-(B_0 - z)q_1$ is decreasing in q_1 because $B_0 > z$ by definition. Hence, $x \mapsto q_1$ is strictly decreasing and continuous as well. This concludes the proof.

We can even go one step further and derive an explicit expression for the derivative of the sovereign debt price with respect to the "autonomous" component of aggregate bailouts. This can be achieved by implicitly differentiating the pricing equation (15), treating q_1 as a function of x. Differentiating on both sides yields

$$R \frac{\partial q_1(x)}{\partial x} = -f\left(B_0 - z + \frac{x}{q_1} \right) \times \frac{q_1 - x \frac{\partial q_1(x)}{\partial x}}{q_1^2}$$

and, solving for the term of interest,

$$\frac{\partial q_1(x)}{\partial x} = \frac{-f\left(B_0 - z + \frac{x}{q_1} \right) \frac{1}{q_1}}{R - f\left(B_0 - z + \frac{x}{q_1} \right) \frac{x}{q_1^2}}$$

Finally, using the Pareto functional form $f(\tau) = \frac{\alpha B_0^\alpha}{\tau^{\alpha+1}}$, we get

$$\frac{\partial q_1(x)}{\partial x} = \frac{\frac{\alpha B_0^\alpha}{\left(B_0 - z + \frac{x}{q_1}\right)^{\alpha+1}} \frac{1}{q_1}}{\frac{\alpha B_0^\alpha}{\left(B_0 - z + \frac{x}{q_1}\right)^{\alpha+1}} \frac{x}{q_1^2} - R} < 0,$$

where the inequality holds because we have shown above that q_1 is decreasing in x. ■

Proof of Proposition 1

The proof proceeds by contradiction. Suppose that $\hat{w}^B < \tilde{w}^B$. Since the parameters (α, γ, B_0) are the same for both financial systems, from equation (7) and Lemma 1 this is true if and only if

$$\hat{T}(\hat{q}_1^B) < \tilde{T}(\tilde{q}_1^B) \quad \text{and} \quad \hat{q}_1^B > \tilde{q}_1^B.$$

Hence, we must have

$$\frac{1 - F\left(B_0 + \frac{\hat{T}(q)}{q}\right)}{R} > \frac{1 - F\left(B_0 + \frac{\tilde{T}(q)}{q}\right)}{R} \quad \forall q \in (0, R^{-1})$$

which is true iff

$$\hat{T}(q) = \sum_{i=1}^{n} \hat{t}^i(q) < \sum_{i=1}^{n} \tilde{t}^i(q) = \tilde{T}(q) \quad \forall q \in (0, R^{-1}).$$

Since $\hat{t}^i(q) = \tilde{t}^i(q)\ \forall q \forall i : \hat{b}^i = \tilde{b}^i$, we must have that

$$\exists i : \left(\hat{b}^i > \tilde{b}^i\right) \wedge \left(\hat{t}^i(q) < \tilde{t}^i(q)\ \forall q \in (0, R^{-1})\right). \tag{17}$$

Using the bailout definition in equation (2) and the second part of the proposition, the second part of (17) reads

$$\left(L^i + d^i \underbrace{-\tilde{c}^i + \frac{\hat{b}^i - \tilde{b}^i}{R}}_{-\hat{c}^i} - q\hat{b}^i - (\pi L)^i\right)^+ < \left(L^i + d^i - \tilde{c}^i - q\tilde{b}^i - (\pi L)^i\right)^+ \quad \forall q \in (0, R^{-1}).$$

Both quantities are bounded from below by zero, so the strict inequality requires that the right-hand side is positive and we can drop the $(\cdot)^+$ operator. The statement thus simplifies to

$$\frac{\hat{b}^i - \tilde{b}^i}{R} < q(\hat{b}^i - \tilde{b}^i) \quad \forall q \in (0, R^{-1}).$$

Finally, since $(\hat{b}^i - \tilde{b}^i) > 0$ from (17), we get

$$q > R^{-1} \quad \forall q \in (0, R^{-1}),$$

a contradiction. We have thus shown that $\hat{w}^B \geq \tilde{w}^B$. From (6) it is easy to see that $\hat{w}(\bar{p}) = \tilde{w}(\bar{p})$ because with $q_1^{NB} = R^{-1}$,

$$\hat{\chi}_1^i = \left(L^i + d^i - \tilde{c}^i + \frac{\hat{b}^i - \tilde{b}^i}{R} - \frac{\hat{b}^i}{R} - (\pi \bar{p})^i \right)^+ = \tilde{\chi}_1^i \quad \forall i \in N$$

Following Definition 3, $\hat{w}^B \geq \tilde{w}^B$ and $\hat{w}(\bar{p}) = \tilde{w}(\bar{p})$ complete the proof of $\hat{\Psi} \subseteq \tilde{\Psi}$. ■

Proof of Proposition 2

The proof proceeds in four steps: First, we argue that there exists a tighter upper bound on the equilibrium sovereign debt price than R^{-1} and that for the financial system $(L, \pi, \tilde{c}, d, \tilde{b})$ this bound is given by $\tilde{q}_1^{max}$ used in the proposition. Second, we demonstrate that banks in $(L, \pi, \tilde{c}, d, \tilde{b})$ which have a positive shortfall at $\tilde{q}_1^{max}$ need a bailout at the equilibrium $\tilde{q}_1$. Third, we show that at the equilibrium $\tilde{q}_1$ the net effect of moving from $(L, \pi, \tilde{c}, d, \tilde{b})$ to $(L, \pi, \hat{c}, d, \hat{b})$ on aggregate bailout expenditures is negative. Finally we prove that this implies $\hat{T}(\hat{q}_1) < \tilde{T}(\tilde{q}_1)$ and thus $\hat{q}_1 > \tilde{q}_1$ which directly yields $\hat{w}^B < \tilde{w}^B$.

First note that because of the inverse relationship of the equilibrium price q_1^B and equilibrium aggregate bailouts $T(q_1^B)$ in Lemma 1, the upper bound on the equilibrium price q_1^{max} is associated with the lower bound on aggregate bailouts. Moreover, the assumption that $\mathcal{F}$ is non-empty and Assumption 1 guarantee that $q_1^B < R^{-1}$. In general, equation (2) thus implies that for all $i \in N$

$$t^i(q_1^B) = \left(L^i + d^i - c^i - q_1^B b^i - (\pi L)^i \right)^+ \geq \left(L^i + d^i - c^i - b^i R^{-1} - (\pi L)^i \right)^+ = \chi_0^i, \tag{18}$$

so the natural candidate for a lower bound on aggregate bailouts is $\sum_{i=1}^n \chi_0^i$. Note that by definition the right-hand side is positive if and only if $i \in \mathcal{F}$. Therefore, (18) holds with equality for every $i \in N$ if and only if two conditions hold:

1. $b^i = 0 \; \forall i \in \mathcal{F}$

2. $L^i + d^i - c^i - q_1^{max} b^i - (\pi L)^i \le 0 \quad \forall i \in \mathcal{B}$

The first condition ensures that the equality holds for all i for which the two sides of (18) are positive. The second part (already using the implicitly characterized q_1^{max}) ensures that it holds for all banks with $b^i > 0$ for which the right-hand side is zero.

We have thus established the existence of an exogenous minimum level of aggregate equilibrium bailouts, namely $\sum_{i=1}^n \chi_0^i$. From the pricing equation (4) it is now straightforward to characterize the corresponding maximum sovereign debt price in $(L, \pi, \tilde{c}, d, \tilde{b})$ as the fixed-point solution to

$$\tilde{q}_1^{max} = \frac{1 - F\left(\frac{\sum_{i \in \mathcal{F}} \chi_0^i}{\tilde{q}_1^{max}}\right)}{R},$$

as in the proposition. Note that even though $\tilde{q}_1^{max}$ cannot be stated explicitly, it only depends on exogenous variables.

Next, let us denote the set of banks from part 1. of the proposition by $Y \equiv \{i : V_0^i / \tilde{b}^i < R^{-1} - \tilde{q}_1^{max}\}$. Their defining property says that for all $i \in Y$

$$\left(\tilde{c}^i + \tilde{b}^i R^{-1} + (\pi L)^i - L^i - d^i\right)^+ < \tilde{b}^i R^{-1} - \tilde{b}^i \tilde{q}_1^{max}$$

which, after rearranging and combining with (2), implies that

$$\tilde{t}^i(\tilde{q}_1^{max}) = \left(L^i + d^i - \tilde{c}^i - \tilde{q}_1^{max} \tilde{b}^i - (\pi L)^i\right)^+ > 0 \quad \forall i \in Y.$$

As we have shown, $\tilde{q}_1^{max}$ is the maximum possible equilibrium price, so in equilibrium it must be also true that $\tilde{t}^i(\tilde{q}_1) > 0 \; \forall i \in Y$.

Next we will show that $\hat{T}(\tilde{q}_1) < \tilde{T}(\tilde{q}_1)$ which is equivalent to $\sum_{i \in N} \hat{t}^i(\tilde{q}_1) - \tilde{t}^i(\tilde{q}_1) < 0$. First note from the definition of bailout transfers (2) that $\hat{t}^i(\tilde{q}_1) = \tilde{t}^i(\tilde{q}_1) \; \forall i \in X \equiv \{i : \tilde{b}^i = \hat{b}^i\}$. In contrast, for banks with $\hat{b}^i < \tilde{b}^i$ (namely $i \in Y$ according to the proposition), we claim that

$$\hat{t}^i(\tilde{q}_1) = L^i + d^i - (\pi L)^i - \tilde{q}_1 \hat{b}^i \underbrace{- \tilde{c}^i - \frac{\tilde{b}^i - \hat{b}^i}{R}}_{-\hat{c}^i} < L^i + d^i - (\pi L)^i - \tilde{c}^i - \tilde{q}_1 \tilde{b}^i = \tilde{t}^i(\tilde{q}_1),$$

where we can drop the $(\cdot)^+$ operators because we have shown above that the right-hand side is positive. Equivalently, the claim can be written as

$$\hat{t}^i(\tilde{q}_1) - \tilde{t}^i(\tilde{q}_1) = (\tilde{q}_1 - R^{-1})(\tilde{b}^i - \hat{b}^i) < 0$$

which is true because $(\tilde{b}^i - \hat{b}^i) > 0 \; \forall i \in Y$ and $\tilde{q}_1 < R^{-1}$. Finally, there are banks $i \in Z \equiv \{i : \hat{b}^i > \tilde{b}^i\}$ for which $\hat{c}^i = \tilde{c}^i - \frac{\hat{b}^i - \tilde{b}^i}{R}$ because initial shortfalls χ_0 are identical in both systems. For these $i \in Z$ we have

$$\hat{t}^i(\tilde{q}_1) = \left(L^i + d^i - (\pi L)^i - \tilde{q}_1 \hat{b}^i \underbrace{-\tilde{c}^i + \frac{\hat{b}^i - \tilde{b}^i}{R}}_{-\hat{c}^i} \right)^+ \geq \left(L^i + d^i - (\pi L)^i - \tilde{c}^i - \tilde{q}_1 \tilde{b}^i \right)^+ = \tilde{t}^i(\tilde{q}_1) \tag{19}$$

since the argument of the $(\cdot)^+$ operator on the left-hand side is strictly larger than that of the right-hand side. There are up to three disjoint subsets of banks to be distinguished: First, for banks where both sides of (19) are zero (call the set Z_1), $\hat{t}^i(\tilde{q}_1) - \tilde{t}^i(\tilde{q}_1) = 0$ is obvious. Second, for banks where the right-hand side is zero and the left-hand side positive (Z_2), we have

$$\hat{t}^i(\tilde{q}_1) - \tilde{t}^i(\tilde{q}_1) = L^i + d^i - (\pi L)^i - \tilde{q}_1 \hat{b}^i - \tilde{c}^i + \frac{\hat{b}^i - \tilde{b}^i}{R} > 0.$$

And third, for banks where both sides are positive (Z_3), we have

$$\hat{t}^i(\tilde{q}_1) - \tilde{t}^i(\tilde{q}_1) = (R^{-1} - \tilde{q}_1)(\hat{b}^i - \tilde{b}^i) > 0,$$

where the inequality follows because both terms are positive for $i \in Z_3$. Now let $\Delta^i \equiv \hat{t}^i(\tilde{q}_1) - \tilde{t}^i(\tilde{q}_1)$ and sum up over all banks:

$$\begin{aligned}
\sum_{i \in N} \Delta^i &= \sum_{i \in X} \Delta^i + \sum_{i \in Y} \Delta^i + \sum_{i \in Z_1} \Delta^i + \sum_{i \in Z_2} \Delta^i + \sum_{i \in Z_3} \Delta^i \\
&= \sum_{i \in Y} \Delta^i + \sum_{i \in Z_2} \Delta^i + \sum_{i \in Z_3} \Delta^i \\
&= (R^{-1} - \tilde{q}_1) \sum_{i \in Y} (\hat{b}^i - \tilde{b}^i) + \sum_{i \in Z_2} \Delta^i + (R^{-1} - \tilde{q}_1) \sum_{i \in Z_3} (\hat{b}^i - \tilde{b}^i) \\
&= (R^{-1} - \tilde{q}_1) \Big\{ \sum_{i \in Y} (\hat{b}^i - \tilde{b}^i) + \sum_{i \in Z_3} (\hat{b}^i - \tilde{b}^i) \Big\} + \sum_{i \in Z_2} L^i + d^i - (\pi L)^i - \tilde{q}_1 \hat{b}^i - \tilde{c}^i + \frac{\hat{b}^i - \tilde{b}^i}{R} \\
&= (R^{-1} - \tilde{q}_1) \Big\{ \sum_{i \in Y} (\hat{b}^i - \tilde{b}^i) + \sum_{i \in Z_3} (\hat{b}^i - \tilde{b}^i) + \sum_{i \in Z_2} \hat{b}^i \Big\} + \sum_{i \in Z_2} L^i + d^i - (\pi L)^i - \tilde{c}^i - \tilde{b}^i R^{-1},
\end{aligned}$$

where we have substituted the expressions for Δ^i derived above for all subsets X, Y, Z. By adding and subtracting $(R^{-1} - \tilde{q}_1)\sum_{i\in Z_2}\tilde{b}^i$, this can be rewritten as

$$\begin{aligned}\sum_{i\in N}\Delta^i = (R^{-1}-\tilde{q}_1)\Big\{ &\overbrace{\sum_{i\in Y}(\hat{b}^i - \tilde{b}^i)}^{<0} + \overbrace{\sum_{i\in Z_2}(\hat{b}^i-\tilde{b}^i)}^{\geq 0} + \overbrace{\sum_{i\in Z_3}(\hat{b}^i-\tilde{b}^i)}^{\geq 0}\Big\} \\ &+ \underbrace{\sum_{i\in Z_2} L^i + d^i - (\pi L)^i - \tilde{c}^i - \tilde{q}_1\tilde{b}^i}_{\leq 0}\end{aligned} \tag{20}$$

Notice that part 2. of the proposition says that $\sum_{i\in N}(\hat{b}^i - \tilde{b}^i) = 0$ which is equivalent to $\sum_{i\in Y\cup Z}(\hat{b}^i - \tilde{b}^i) = 0$.[25] Hence, if the set Z_1 was empty and therefore $N = X \cup Y \cup Z_2 \cup Z_3$, the term in curly brackets would be zero and

$$\sum_{i\in N}\Delta^i = \sum_{i\in Z_2} L^i + d^i - (\pi L)^i - \tilde{c}^i - \tilde{q}_1\tilde{b}^i \leq 0,$$

where the inequality follows from the definition of the set Z_2 and is strict except for the knife-edge case in which every Z_2-bank was *just* solvent in the tilde-equilibrium (i.e., assets exactly matched liabilities). If, in addition to Z_1, the set Z_2 was also empty, the last term in (20) would disappear and we would have $\sum_{i\in N}\Delta^i = 0$, so there would be no change in bailouts ($\hat{T}(\tilde{q}_1) = \tilde{T}(\tilde{q}_1)$) and the two equilibria would coincide, i.e., $\hat{T}(\hat{q}_1) = \tilde{T}(\tilde{q}_1)$ and $\hat{q}_1 = \tilde{q}_1$. However, as soon as there is at least one Z_1-bank (or a Z_2-bank with positive equity in $(L, \pi, \tilde{c}, d, \tilde{b})$), equation (20) yields $\hat{T}(\tilde{q}_1) < \tilde{T}(\tilde{q}_1)$, as desired.

Finally, using the right-hand side of the pricing equation (4), notice that the previous result implies

$$\frac{1 - F(B_0 + \frac{\hat{T}(\tilde{q}_1)}{\tilde{q}_1})}{R} > \frac{1 - F(B_0 + \frac{\tilde{T}(\tilde{q}_1)}{\tilde{q}_1})}{R} = \tilde{q}_1$$

As we show in Appendix A, under the Pareto assumption this expression increases monotonically in q_1 and is strictly concave everywhere, including at $\tilde{q}_1$. Therefore, if $\tilde{q}_1$ is the equilibrium price in $(L, \pi, \tilde{c}, d, \tilde{b})$, the unique fixed point $\hat{q}_1$ in $(L, \pi, \hat{c}, d, \hat{b})$ must be larger ($\hat{q}_1 > \tilde{q}_1$). As a consequence, since we have established above that $\hat{T}(\tilde{q}_1) < \tilde{T}(\tilde{q}_1)$ is true, *a fortiori* $\hat{T}(\hat{q}_1) < \tilde{T}(\tilde{q}_1)$ also holds because $\hat{T}(\cdot)$ is a decreasing function.

---

[25] Recall that the premise of the proposition is that sovereign debt is redistributed from banks in Y to banks in Z, so that $\sum_{i\in Y}(\hat{b}^i - \tilde{b}^i) = -\sum_{i\in Z}(\hat{b}^i - \tilde{b}^i)$.

It now follows directly from (7) that $\hat{w}^B < \tilde{w}^B$. Moreover, as in Proposition 1 we have $\hat{w}(\bar{p}) = \tilde{w}(\bar{p})$ because with $q_1^{NB} = R^{-1}$,

$$\hat{\chi}_1^i = \left(L^i + d^i - \tilde{c}^i + \frac{\hat{b}^i - \tilde{b}^i}{R} - \frac{\hat{b}^i}{R} - (\pi\bar{p})^i \right)^+ = \tilde{\chi}_1^i \quad \forall i \in N$$

Following Definition 3, $\hat{w}^B < \tilde{w}^B$ and $\hat{w}(\bar{p}) = \tilde{w}(\bar{p})$ complete the proof of $\hat{\Psi} \supset \tilde{\Psi}$. ■

Proof of Proposition 3

The proof proceeds in three steps. In the first step we show that $\tilde{\Delta}^i w > \tilde{\Delta}^j w$; in the second that $\hat{\Delta}^i w > \hat{\Delta}^j w$; and in the last step that $\hat{\Delta}^i w - \hat{\Delta}^j w > \tilde{\Delta}^i w - \tilde{\Delta}^j w$.

Using the full expression for the net welfare effect from Lemma 2, we want to show that

$$\beta \sum_{k \in \mathcal{I}} \chi_1^k \times C^k - \gamma \tilde{\Delta}^i (B_1 P(Default)) > \beta \sum_{l \in \mathcal{J}} \chi_1^l \times C^l - \gamma \tilde{\Delta}^j (B_1 P(Default))$$

Note that by assumption $\chi_1^i = \chi_1^j$ and hence $\tilde{\Delta}^i(B_1 P(Default)) = \tilde{\Delta}^j(B_1 P(Default))$, so the expression becomes

$$\beta \sum_{k \in \mathcal{I}} \chi_1^k \times C^k > \beta \sum_{l \in \mathcal{J}} \chi_1^l \times C^l.$$

Since $\mathcal{I} \setminus i = \mathcal{J} \setminus j$ by assumption, this simplifies to

$$\beta \chi_1^i C^i > \beta \chi_1^j C^j$$

which is true because $C^i > C^j$.

We now turn to the second step. Using Lemma 2 and the same facts as above (i.e. identical gamma-components and $\mathcal{I} \setminus i = \mathcal{J} \setminus j$) the expression we want to prove becomes

$$\beta \chi_1^i C^i + \beta \hat{\Delta}^i q_1 b' C^j > \beta \chi_1^j C^j + \beta \hat{\Delta}^j q_1 b' C^i$$

Identical shortfalls $\chi_1^i = \chi_1^j$ together with identical sovereign exposure $\hat{b}^i = \hat{b}^j = b' > 0$ imply an identical price impact $\hat{\Delta}^i q_1 = \hat{\Delta}^j q_1$, so we have

$$\chi_1^i (C^i - C^j) > \hat{\Delta}^i q_1 b' (C^i - C^j)$$

The left-hand side is clearly positive because $C^i > C^j$ whereas the right-hand side is negative because $\hat{\Delta}^i q_1 < 0$, so we have the desired result.

Finally, it remains to be shown that $\hat{\Delta}^i w - \hat{\Delta}^j w > \tilde{\Delta}^i w - \tilde{\Delta}^j w$. Again, using Lemma 2, identical gamma-components, and $\mathcal{I} \setminus i = \mathcal{J} \setminus j$ the expression becomes

$$\beta \chi_1^i (C^i - C^j) - \beta \hat{\Delta}^i q_1 b' (C^i - C^j) > \beta \chi_1^i (C^i - C^j).$$

That simplifies to

$$\beta \hat{\Delta}^i q_1 b' (C^i - C^j) < 0$$

which is true because $\hat{\Delta}^i q_1 < 0$. That concludes the proof. ■

Proof of Proposition 4

To prove the proposition it is sufficient to show that $\Delta^i w - \Delta^j w$ is monotonically increasing in C and that $\exists \tilde{C}$ such that $\Delta^i w = \Delta^j w$.

From Lemma 2 we can write $\Delta^i w - \Delta^j w$ as

$$\beta \left(\sum_{h \in \mathcal{I}} \chi_1^h C^h - \sum_{k \in \mathcal{J}} \chi_1^k C^k \right) + \beta \left(\sum_{h \in \mathcal{S}^c \setminus \mathcal{I}} \Delta^i q_1 b^h C^h - \sum_{k \in \mathcal{S}^c \setminus \mathcal{J}} \Delta^j q_1 b^k C^k \right) - \gamma \Gamma,$$

where $\Gamma = \Delta^i (B_1 P(Default)) - \Delta^j (B_1 P(Default)) > 0$ because $b^i > b^j$ implies that $t^i > t^j$ due to doom loop multiplier and $\mathcal{I}$ and $\mathcal{J}$ denote the sets of banks that become solvent through a bailout of bank i and j, respectively. Now note that because $\mathcal{I} \setminus i = \mathcal{J} \setminus j$, $\chi_1^i = \chi_1^j = \chi_1$ and $C^i = C^j = C$, the benefit components of $\Delta^i w$ and $\Delta^j w$ are identical, so the first term collapses to zero.

The two sums in the second term in parentheses contain common terms, namely for all banks that still default after saving i or j, excluding i and j themselves. Therefore, the previous expression can be rewritten as

$$\beta \left(\Delta^i q_1 b^j C - \Delta^j q_1 b^i C + (\Delta^i q_1 - \Delta^j q_1) \sum_{l \in \mathcal{S}^c \setminus (\mathcal{I} \cup \mathcal{J})} b^l C^l \right) - \gamma \Gamma,$$

where the set $\mathcal{S}^c \setminus (\mathcal{I} \cup \mathcal{J})$ consists of the banks (different from i and j) that still default after saving either i or j. We can now isolate the term multiplying C, namely $\beta(\Delta^i q_1 b^j - \Delta^j q_1 b^i)$. The premise of the proposition that $\frac{b^i}{b^j} > \frac{\Delta^i q_1}{\Delta^j q_1} > 1$ ensures that this coefficient is positive, so we have established that $\Delta^i w - \Delta^j w$ is monotonically increasing in C.

The existence of $\tilde{C}$ such that $\Delta^i w = \Delta^j w$ can easily be shown by setting the previous expression equal to zero. We get

$$\beta\tilde{C}(\Delta^i q_1 b^j - \Delta^j q_1 b^i) = \gamma\Gamma - \beta(\Delta^i q_1 - \Delta^j q_1) \sum_{l\in\mathcal{S}^c\setminus(\mathcal{I}\cup\mathcal{J})} b^l C^l$$

and finally solve for the threshold level of centrality

$$\tilde{C} = (\underbrace{\Delta^i q_1 b^j - \Delta^j q_1 b^i}_{>0})^{-1} \left(\frac{\gamma}{\beta}\Gamma - (\underbrace{\Delta^i q_1 - \Delta^j q_1}_{<0}) \sum_{l\in\mathcal{S}^c\setminus(\mathcal{I}\cup\mathcal{J})} b^l C^l\right) > 0$$

That concludes the proof. ■

Proof of Proposition 5

To prove the proposition it is sufficient to show that $(\hat{\Delta}^i w - \hat{\Delta}^j w) - (\tilde{\Delta}^i w - \tilde{\Delta}^j w)$ is monotonically increasing in C^i and that there exists a $C^*(j)$ such that $\hat{\Delta}^i w - \hat{\Delta}^j w = \tilde{\Delta}^i w - \tilde{\Delta}^j w$.

From Lemma 2 we can rewrite the difference as

$$\begin{aligned}
\beta\Big[\sum_{k\in\mathcal{I}}\chi_1^k C^k - \sum_{l\in\mathcal{J}}\chi_1^l C^l\Big] + \beta\Big[\sum_{k\in\mathcal{S}^c\setminus\mathcal{I}}\hat{\Delta}^i q_1 \hat{b}^k C^k - \sum_{l\in\mathcal{S}^c\setminus\mathcal{J}}\hat{\Delta}^j q_1 \hat{b}^l C^l\Big] \\
-\gamma\Big[\hat{\Delta}^i(B_1 P(Default)) - \hat{\Delta}^j(B_1 P(Default))\Big] \\
-\beta\Big[\sum_{k\in\mathcal{I}}\chi_1^k C^k - \sum_{l\in\mathcal{J}}\chi_1^l C^l\Big] - \beta\Big[\sum_{k\in\mathcal{S}^c\setminus\mathcal{I}}\tilde{\Delta}^i q_1 \tilde{b}^k C^k - \sum_{l\in\mathcal{S}^c\setminus\mathcal{J}}\tilde{\Delta}^j q_1 \tilde{b}^l C^l\Big] \\
+\gamma\Big[\tilde{\Delta}^i(B_1 P(Default)) - \tilde{\Delta}^j(B_1 P(Default))\Big]
\end{aligned}$$

Now note that the benefit terms (the first and the fourth term) cancel each other out exactly. Moreover, the γ-components of bailing out banks $j \neq i$ is the same in both tilde- and hat-system (because $\hat{b}^j = \tilde{b}^j$ and $\hat{\chi}_1^j = \tilde{\chi}_1^j$), so $\hat{\Delta}^j(B_1 P(Default)) = \tilde{\Delta}^j(B_1 P(Default))$ and the expression simplifies to

$$\begin{aligned}
\beta\Bigg[\left(\sum_{k\in\mathcal{S}^c\setminus\mathcal{I}}\hat{\Delta}^i q_1 \hat{b}^k C^k - \sum_{k\in\mathcal{S}^c\setminus\mathcal{I}}\tilde{\Delta}^i q_1 \tilde{b}^k C^k\right) - \left(\sum_{l\in\mathcal{S}^c\setminus\mathcal{J}}\hat{\Delta}^j q_1 \hat{b}^l C^l - \sum_{l\in\mathcal{S}^c\setminus\mathcal{J}}\tilde{\Delta}^j q_1 \tilde{b}^l C^l\right)\Bigg] \\
-\gamma\left(\hat{\Delta}^i(B_1 P(Default)) - \tilde{\Delta}^i(B_1 P(Default))\right)
\end{aligned}$$

For the sake of notation, denote $\Gamma \equiv \hat{\Delta}^i(B_1 P(Default)) - \tilde{\Delta}^i(B_1 P(Default)) > 0$ because $\hat{b}^i > \tilde{b}^i$. We can then rewrite the previous expression as follows:

$$\beta\left(\sum_{k\in\mathcal{S}^c\setminus\mathcal{I}}(\hat{\Delta}^i q_1\hat{b}^k - \tilde{\Delta}^i q_1\tilde{b}^k)\times C^k - \sum_{l\in\mathcal{S}^c\setminus\mathcal{J}}(\hat{\Delta}^j q_1\hat{b}^l - \tilde{\Delta}^j q_1\tilde{b}^l)\times C^l\right) - \gamma\Gamma$$

Now note that the proposition says that $\forall k \in \mathcal{S}^c \setminus \mathcal{I}$ we have $\hat{b}^k = \tilde{b}^k$ (first term in round brackets). Then we can simplify the expression to

$$\beta\left(\sum_{k\in\mathcal{S}^c\setminus\mathcal{I}}(\hat{\Delta}^i q_1 - \tilde{\Delta}^i q_1)\hat{b}^k C^k - \sum_{l\in\mathcal{S}^c\setminus\mathcal{J}}\hat{\Delta}^j q_1(\hat{b}^l - \tilde{b}^l)\times C^l\right) - \gamma\Gamma$$

where we also use the fact that $\hat{\Delta}^l q_1 = \tilde{\Delta}^l q_1 \ \forall l \neq j$. Finally, since $\hat{b}^l = \tilde{b}^l \ \forall l \neq i$, the previous expression reduces to

$$\beta\left(\left[(\hat{\Delta}^i q_1 - \tilde{\Delta}^i q_1)\sum_{k\in\mathcal{S}^c\setminus\mathcal{I}}\hat{b}^k C^k\right] - \hat{\Delta}^j q_1(\hat{b}^i - \tilde{b}^i)C^i\right) - \gamma\Gamma$$

Now it becomes clear that the expression is monotonically increasing in C^i, since $\hat{\Delta}^j q_1 < 0$ and $\hat{b}^i > \tilde{b}^i$. In a last step, we can solve for the threshold $C^*(j)$ that sets the above expression equal to zero and thereby prove its existence:

$$C^*(j) = (\underbrace{|\hat{\Delta}^j q_1|(\hat{b}^i - \tilde{b}^i)}_{>0})^{-1}\left(\frac{\gamma}{\beta}\Gamma - (\underbrace{\hat{\Delta}^i q_1 - \tilde{\Delta}^i q_1}_{<0})\sum_{k\in\mathcal{S}^c\setminus\mathcal{I}}\hat{b}^k C^k\right) > 0$$

That concludes the proof. ■

C Spectral Radius Assumption

The spectral radius of the matrix $(1+\beta)\Pi'_{\mathcal{D}}$ is given by its largest eigenvalue (in absolute value). In the following we compute the spectral radius first for $|\mathcal{D}| = 2$ and then for $|\mathcal{D}| = 3$ and derive conditions such that it is smaller than unity.

If the default set consists of two banks (and hence $\Pi'_{\mathcal{D}}$ is a 2×2 matrix), to compute the eigenvalues λ_1, λ_2 we solve the equation

$$\left|(1+\beta)\Pi'_{\mathcal{D}} - \lambda\mathbb{I}_2\right| = 0$$

and obtain $\lambda_{1/2} = \pm(1+\beta)\sqrt{\pi^{12}\pi^{21}}$.[26] Hence, the spectral radius is less than unity if and only if

$$\pi^{12}\pi^{21} < (1+\beta)^{-2}$$

If the default set consists of three banks (and hence $\Pi'_{\mathcal{D}}$ is a 3×3 matrix) we proceed in the same way. We set the determinant of $(1+\beta)\Pi'_{\mathcal{D}} - \lambda \mathbb{I}_3$ equal to zero and obtain the characteristic polynomial

$$\lambda^3 - \underbrace{(1+\beta)^2\left(\pi^{13}\pi^{31} + \pi^{23}\pi^{32} + \pi^{21}\pi^{12}\right)}_{\equiv p}\lambda - \underbrace{(1+\beta)^3\left(\pi^{13}\pi^{32}\pi^{21} + \pi^{12}\pi^{23}\pi^{31}\right)}_{\equiv q} = 0 \tag{21}$$

which is a depressed cubic. Therefore, one of the solutions is guaranteed to be real, can be obtained by applying the Cardano formula and is given by

$$\lambda_1 = \sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} \tag{22}$$

The remaining two roots λ_2, λ_3 are

$$\lambda_2 = \sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} \times \left(-\frac{1}{2} + \frac{i\sqrt{3}}{2}\right) + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} \times \left(-\frac{1}{2} - \frac{i\sqrt{3}}{2}\right)$$

and

$$\lambda_3 = \sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} \times \left(-\frac{1}{2} - \frac{i\sqrt{3}}{2}\right) + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} \times \left(-\frac{1}{2} + \frac{i\sqrt{3}}{2}\right)$$

These expressions can be rewritten as

$$\lambda_2 = -\frac{1}{2}\lambda_1 + \frac{i\sqrt{3}}{2}\underbrace{\left(\sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} - \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}}\right)}_{\equiv\tilde{\lambda}_2}$$

and

$$\lambda_3 = -\frac{1}{2}\lambda_1 + \frac{i\sqrt{3}}{2}\underbrace{\left(\sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} - \sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}}\right)}_{\equiv\tilde{\lambda}_3}$$

---

[26] Recall that $\pi^{ii} = 0 \; \forall i \in N$ by construction of Π. The submatrix $\Pi_{\mathcal{D}}$ inherits this property.

Here we can see that iff $4p^3 < 27q^2$, the argument of the square roots is positive and hence λ_2, λ_3 are complex numbers. To find out the largest eigenvalue, we thus have to compare $|\lambda_1|$ to the *modulus* of the two complex roots. First note that $(\tilde{\lambda}_2)^2 = (\tilde{\lambda}_3)^2$ and hence

$$|\lambda_2| = \sqrt{\frac{1}{4}(\lambda_1)^2 + \frac{3}{4}(\tilde{\lambda}_2)^2} = |\lambda_3|,$$

so the two complex roots have the same absolute value. Now note that $\lambda_1 > \tilde{\lambda}_2 > 0$ and hence also $(\lambda_1)^2 > (\tilde{\lambda}_2)^2$, so we can write

$$|\lambda_2| = \sqrt{\frac{1}{4}(\lambda_1)^2 + \frac{3}{4}(\tilde{\lambda}_2)^2} < \sqrt{\frac{1}{4}(\lambda_1)^2 + \frac{3}{4}(\lambda_1)^2} = \sqrt{(\lambda_1)^2} = |\lambda_1|,$$

so we have established that the real root $|\lambda_1|$ is the spectral radius iff $4p^3 < 27q^2$. In that case the condition we are looking for reads

$$\left| \sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} \right| < 1.$$

If instead $4p^3 > 27q^2$, the argument of the square roots is negative and it can be shown that all three roots $\lambda_1, \lambda_2, \lambda_3$ are real in this case.

In the knife-edge case of $4p^3 = 27q^2$, the square roots are equal to zero and we have

$$\lambda_1 = 2\sqrt[3]{\frac{q}{2}}$$
$$\lambda_2 = \lambda_3 = -\frac{\lambda_1}{2},$$

so again $|\lambda_1|$ is the largest eigenvalue and our condition $|\lambda_1| < 1$ becomes

$$q = (1+\beta)^3 \left(\pi^{13}\pi^{32}\pi^{21} + \pi^{12}\pi^{23}\pi^{31} \right) < \frac{1}{4}$$

We conclude this section with two stylized examples. First suppose that the interbank network forms a *ring* of length $n = 3$ in which all three banks default, i.e., $\pi^{ij} = 1$ for $j = i+1$, modulo n, and $\pi^{ij} = 0$ otherwise. In that case note from (21) that $p = 0$ and $q = (1+\beta)^3$, so we are in the first case with $4p^3 < 27q^2$ and the spectral radius is given by the unique real root in (22). The condition for node depth to be well defined reads

$$|\lambda_1| = |\sqrt[3]{q}| = 1 + \beta < 1$$

which is a contradiction because $\beta > 0$ by assumption. Hence, if the network is a ring and all three banks default, node depth cannot be computed using the matrix inversion formula.

As a second example, suppose that the interbank network is *complete*, that is $\pi^{ij} = \frac{1}{n-1} \forall i, j, i \neq j$. For ease of notation, let $k = n - 1$ so that from (21) we know that

$$p = (1+\beta)^2 \left(k^{-2} + k^{-2} + k^{-2}\right) = \frac{3(1+\beta)^2}{k^2}$$

$$q = (1+\beta)^3 \left(k^{-3} + k^{-3}\right) = \frac{2(1+\beta)^3}{k^3}$$

It is easy to verify that in this case we get exactly $4p^3 = 27q^2$, so the condition for C to be well defined is given by $q < \frac{1}{4}$ as shown above. Plugging in the expression for q yields

$$k > 2(1+\beta)$$

Hence, node depth in a default set of three banks is well defined iff the total number of banks in the network n is larger than $1 + 2(1+\beta)$. The critical number of banks is increasing in the bankruptcy cost parameter β.

D Algorithm to Find Equilibria with Optimal Bailouts

First note that any chosen subset of surviving banks $\mathcal{S}$ induces unique transfers $\left(t^i(\bar{p}(\mathcal{S}), q_1)\right)_{i \in \mathcal{S}}$ according to equation (8). To determine these transfers (and the associated sovereign debt prices) we proceed as follows *for each possible subset* $\mathcal{S} \subseteq N$:

1. Starting with an initial guess of $q_1 = 1/R$, we find the Pareto dominant clearing payment vector $\bar{p}(\mathcal{S})$ as per Definition 1, imposing that all surviving banks repay their liabilities in full, i.e. $\bar{p}^i = L^i \; \forall i \in \mathcal{S}$.

2. Knowing the clearing payment vector $\bar{p}(\mathcal{S})$, we compute the required bailout transfers to $\mathcal{S}$-banks from equation (8), as a function of the current guess for q_1.

3. Aggregate bailouts $T(\bar{p}(\mathcal{S}), q_1) = \sum_{i \in \mathcal{S}} t^i(\bar{p}(\mathcal{S}), q_1)$ pin down the sovereign debt price $q_1 < 1/R$ via equation (4). We then repeat the procedure with the updated guess for q_1 and iterate until convergence.

4. Compute welfare losses $w(\bar{p}(\mathcal{S}))$ using equation (9).

The optimal set $\mathcal{S}^*$ is the set with the lowest welfare losses.