Systemic Risk and Stability in Financial Networks[*]

Daron Acemoglu[†]   Asuman Ozdaglar[‡]   Alireza Tahbaz-Salehi[§]

This version: January 2013
First version: December 2011

Abstract

We provide a framework for studying the relationship between the financial network architecture and the likelihood of systemic failures due to contagion of counterparty risk. We show that financial contagion exhibits a form of *phase transition* as interbank connections increase: as long as the magnitude and the number of negative shocks affecting financial institutions are sufficiently small, more "complete" interbank claims enhance the stability of the system. However, beyond a certain point, such interconnections start to serve as a mechanism for propagation of shocks and lead to a more fragile financial system. We also show that, under natural contracting assumptions, financial networks that emerge in equilibrium may be socially inefficient due to the presence of a network externality: even though banks take the effects of their lending, risk-taking and failure on their immediate creditors into account, they do not internalize the consequences of their actions on the rest of the network.

Keywords: Contagion, counterparty risk, financial network, systemic risk.
JEL Classification: G01, D85.

1 Introduction

Since the global financial crisis of 2008, the view that the architecture of the financial system causes, and shapes the nature of, crises has become conventional wisdom. The intertwined nature of financial markets has not only been offered as an explanation for the spread of risk throughout the system,[1] but also motivated much of the policy actions as the crisis unfolded.[2] Such views have even been incorporated into the new regulatory frameworks developed since.[3] Yet, the specific role played by the financial network structure in creating systemic risk and shaping the fragility of the financial system remains, at best, imperfectly understood. This is not only due to a lack of conclusive empirical evidence on the nature of financial contagion, but also due to the absence of a theoretical framework that can serve to clarify the relevant economic forces.

The current state of uncertainty about the nature and causes of systemic risk is reflected in potentially conflicting views on the relationship between the structure of the financial network and the extent of financial contagion. Pioneering works by Allen and Gale (2000) and Freixas, Parigi, and Rochet (2000) suggested that a more equal distribution of interbank claims enhances the resilience of the system to the insolvency of any individual bank.[4] Allen and Gale, for example, argue that in a more densely interconnected financial network, the losses of a distressed bank are divided among more creditors, reducing the impact of negative shocks to individual institutions on the rest of the system. In contrast to this view, however, others have suggested that dense interconnections may function as a destabilizing force, paving the way to systemic failures. For example, Vivier-Lirimont (2006) argues that as the number of a bank's counterparties grows, the likelihood of a systemic collapse increases. This perspective is also shared by Blume et al. (2011) who model interbank contagion as an epidemic.[5]

In view of the conflicting perspectives noted above, this paper presents a unifying, yet simple framework for studying the role of the financial system's architecture in shaping systemic risk. In particular, by focusing on the relationship between the structure of the financial network and the extent of contagion via the so-called domino effects, we clarify the economic forces that may contribute to the propagation of shocks during the times of crises.[6]

---

[1]See, for example, Plosser (2009) and Financial Crisis Inquiry Commission (2011).

[2]For an account of the policy actions during the crisis, see Sorkin (2009). For a general discussion of the crisis and events that followed the collapse of Lehman Brothers, see Gorton (2010). Also see Gorton and Metrick (2011) on the spread of the crisis from subprime housing assets to the repo market; Stulz (2010) on the CDS market; and Longstaff (2010) on cross-market contagion.

[3]For example, the provision on "single counterparty exposure limits" in the Dodd-Frank Wall Street Reform and Consumer Protection Act is meant to prevent the distress at a single institution from infecting the rest of the system by limiting each firm's exposure to any single counterparty (*The Economist*, 05/12/12).

[4]Similar ideas are discussed by Kiyotaki and Moore (1997), who study contagion in credit chains amongst lenders and entrepreneurs.

[5]Along the same lines, Battiston et al. (2012) argue that due to a positive feedback loop resulting from a financial acceleration mechanism, more interconnectivity leads to greater fragility in the financial system. Somewhat relatedly, Billio et al. (2012) show that, over the past decade, financial institutions have become highly interrelated, which has led to an increase in the level of systemic risk in the finance and insurance industries.

[6]The role played by the network architecture crucially depends on the nature of economic interactions between different entities that constitute the network. For example, focusing on the intersectoral input-output linkages in the real economy, Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012) show that in the presence of linear (or log-linear) economic inter-

We focus on an economy consisting of n financial institutions (henceforth banks, for short) that lasts for three dates. In the initial date, banks borrow funds from one another to invest in projects that yield returns both in the intermediate and final dates. The liability structure that emerges from such interbank loans determines the *financial network*, capturing the pairwise counterparty relationships between different institutions. Banks in our model have to make other payments (such as wages or taxes) with claims that are senior to other banks'. Assuming that the returns at the final date are not pledgeable, all debts have to be repaid at the intermediate date. Thus, a bank whose short-term returns are below a certain level may have to liquidate its project prematurely (i.e., before the final date returns are realized). If the revenues from the liquidations are insufficient to pay all its debts, the bank defaults. Depending on the structure of the financial network, this may then trigger a cascade of failures: the default of a bank on its debt may lead to financial distress of its creditor banks, which in turn may default on their own counterparties, and so on.[7]

We first study the extent of financial contagion while taking the structure of the financial network as given. By generalizing the results of Eisenberg and Noe (2001), we show that a mutually consistent collection of repayments on interbank loans always exists and is generically unique. We then show that when the magnitude and the number of negative shocks are below certain thresholds, a result similar to those of Allen and Gale (2000) and Freixas, Parigi, and Rochet (2000) holds: a more equal distribution of interbank obligations (liabilities) leads to a less fragile financial system. In particular, the complete financial network, in which the liabilities of each institution are equally held by all other banks, is the configuration least prone to contagious defaults. At the opposite end of the spectrum, the ring network — a configuration in which all liabilities of a bank are held by a single counterparty — is the most fragile of all financial network structures. The intuition underlying these results is that a more equal distribution of interbank liabilities guarantees that the burden of any potential losses is shared among more counterparties and hence, in the presence of relatively small shocks, the excess liquidity of the non-distressed banks can be efficiently utilized in forestalling further defaults.

One of our main results is that that as the magnitude or the number of negative shocks cross certain thresholds, the types of financial networks that are most prone to contagious failures change dramatically. In particular, more financial interconnections are no longer a guarantee for stability. Rather, in the presence of large shocks, interbank liabilities facilitate financial contagion and create

---

actions, the volatility of aggregate output is independent of the sparseness or denseness of connections. Rather, it depends on the extent of asymmetry in the interconnections of different entities. Acemoglu, Ozdaglar, and Tahbaz-Salehi (2010) show a similar result for the "systemic" event in which aggregate output falls below a certain threshold. Such conclusions, however, do not extend to financial interactions, as the possibility of default (and the presence of debt-like financial instruments or other non-linearities) create a very different set of economic interactions over the network.

[7]As this description clarifies, our main focus is on the spread of counterparty risk, which is a specific type of financial network interaction. At least two other types of network interactions are important in practice: (i) fire sales of some assets by a bank may create distress on other institutions that hold similar assets (represented by the means of the network of common asset holdings); and (ii) withdrawal of liquidity by a bank (for example, by not rolling over a repo agreement or increasing the haircut on a collateral) may lead to a chain reaction spreading over a particular network structure, perhaps related to the network of counterparty relations. In view of the comment in footnote 6, we expect the role of network architecture in these alternative mechanisms, though related, to be different from the one studied in here. The study of these alternative propagation mechanisms is beyond the scope of the current paper.

a more fragile system. Our results show that, in the presence of large shocks, "weakly connected" financial networks — for example, one consisting of a collection of pairwise connected banks with only a minimal amount of shared assets and liabilities with the rest of the system — are significantly less fragile than the more complete networks.[8] The intuition underlying such a sharp "phase transition" is that, with large negative shocks, the excess liquidity of the banking system may no longer be sufficient for absorbing the losses. Under such a scenario, less interbank connections guarantee that the losses are shared with the senior creditors of the distressed banks, and hence, protecting the rest of the system.

Our formal results thus confirm a conjecture by Haldane (2009) who suggested that highly interconnected financial networks may be "robust-yet-fragile" and that they "exhibit a knife-edge or tipping point property", in the sense that "within a certain range, connections serve as shock-absorbers [and] connectivity engenders robustness." However, beyond a certain range, interconnections start to serve as a mechanism for propagation of shocks, in Haldane's words, "the system [flips to] the wrong side of the knife-edge," and fragility prevails. On a broader level, our results highlight that the same features that make a financial system more resilient under certain conditions may function as significant sources of systemic risk and instability under another.

We also endogenize the interbank counterparty relations and use our characterization of financial contagion to investigate the efficiency of equilibrium financial networks. The key endogeneity in our model involves the structure and the terms of bilateral interbank agreements. In particular, we assume that banks lend to one another through debt contracts with contingency covenants, which allow lenders to charge different interest rates depending on the risk-taking behavior of the borrower. The presence of such covenants is both empirically plausible and theoretically central as it guarantees that all bilateral externalities are internalized. Nevertheless, our key result here is that, despite the covenants, the equilibrium financial networks are generally inefficient. Our results thus highlight the presence of a *financial network externality* in the formation of financial networks: even though banks take the effects of their actions on their immediate creditors into account, they fail to internalize the externalities that they impose on the rest of the network — such as on their creditors' creditors and so on. We then illustrate the implications of the financial network externality for the types of inefficiencies that may arise in the formation of financial networks. In particular, we show that (i) banks may "overlend" in equilibrium, creating channels over which idiosyncratic shocks can translate into systemic crises via financial contagion; and (ii) they may lend to the "wrong" set of borrowers from the point of view of the social planner, creating an insufficiently connected financial network that is excessively prone to contagion. We also show that banks' private incentives may lead to the formation of robust-yet-fragile networks of counterparty relations, which are overly susceptible (from the social planner's point of view) to systemic meltdowns with some small probability.

Finally, we relax the assumption of non-pledgeability of the long-term returns and study the na-

---

[8]Such financial networks are somewhat reminiscent of the old-style unit banking system, in which banks within a region are only weakly connected to the rest of the financial network.

ture of contagion within the financial network. With limited pledgeablity of returns, there is room for renegotiating debts that are due in the intermediate date. We show that when the financial network is connected, there is a threshold of financial distress below which the structural details of the network become irrelevant and the excess liquidity within the banking system can be efficiently mobilized to forestall all defaults. However, once the size of the negative shock is sufficiently large, our earlier results become applicable once again and the fragility of the banking system would highly depend on the structure of the financial network. Thus, limited pledgeability leads to another form of phase transition in the financial network: below the critical threshold, the banking system successfully manages aggregate liquidity provision, whereas above that threshold, the local structural properties of the financial network dictate the extent of contagion.

Our paper is part of a recent but growing literature that focuses on the role of the architecture of the financial system as an amplification mechanism. Allen and Gale (2000) and Freixas, Parigi, and Rochet (2000) provided some of the first formal models of contagion over financial networks. Using a multi-region version of Diamond and Dybvig (1983)'s model, Allen and Gale show that the interbank relations that emerge to pool region-specific shocks may at the same time create fragility in response to unanticipated shocks. Dasgupta (2004) studies how the cross-holdings of deposits motivated by imperfectly correlated regional liquidity shocks can lead to contagious breakdowns. More recently, Allen, Babus, and Carletti (2012) show that the pattern of asset commonalities between different banks determines the extent of information contagion and hence, the likelihood of a systemic crisis. Elsinger (2009) and Elliot, Golub, and Jackson (2012) focus on a different mechanism, according to which cross-holdings of different organizations' shares or debts may lead to cascading failures. Also related is the work of Castiglionesi, Feriozzi, and Lorenzoni (2010), who show that a higher degree of financial integration leads to more stable interbank interest rates in normal times, but to larger interest rate spikes during crises.[9] None of the above papers, however, provide a comprehensive analysis of the relationship between the architecture of the financial network and the likelihood of systemic failures due to contagion of counterparty risk.

The paper perhaps most closely related to ours is an independent work by Eboli (2012), who also studies contagion of counterparty risk. He develops a different analysis based on the network flow problems and provides conditions for indeterminacy of the interbank payments due to the cyclical entanglement of assets and liabilities. Though Eboli also studies the extent of contagion in some classes of networks and notes the possibility of phase transitions, his focus and results are quite different from ours.

Our analysis of the financial externality is also related to a smaller literature on the formation of financial networks. Babus (2009) studies a model in which banks form linkages with one another in order to insure against the risk of contagion. She shows that banks can succeed in forming networks

---

[9]Other related contributions include Cifuentes, Ferrucci, and Shin (2005), Leitner (2005), Rotemberg (2011), Zawadowski (2011), Cabrales, Gottardi, and Vega-Redondo (2012), and Caballero and Simsek (2012). For a more detailed discussion of the literature, see the survey by Allen and Babus (2009).

that are highly resilient to the propagation of shocks. Zawadowski (2012), on the other hand, shows that banks may choose not to buy default insurance on their counterparties, even though this may be socially desirable. This differs from our focus, which is to explicitly endogenize the network of interbank liabilities and study the implications of the financial network externality across institutions.

Finally, our paper is related to the literature that emphasizes the possibility of *indirect* spillovers in the financial market. For example, Shleifer and Vishny (1992), Holmström and Tirole (1998), Brunnermeier and Pedersen (2005), Lorenzoni (2008) and Krishnamurthy (2010) study the potential impacts of firm-level distress on market distress or liquidity shortages. However, rather than taking place through *direct* contractual relations as in our paper, the amplification mechanisms studied in these papers work through the endogenous responses of various market participants.[10]

The rest of the paper is organized as follows. Our model is presented in Section 2. Section 3 contains our results on the relationship between the extent of financial contagion and the financial network acrhitecture. In Section 4, we endogenize the interbank counterparty relations and study the financial networks that arise in equilibrium. In Section 5, we show how the possibility and extent of renegotiations affect the robustness of the financial system to distress. Section 6 concludes. All proofs are presented in the Appendix.

2 Model

2.1 Agents and Technology

Consider a single good economy, consisting of n risk-neutral banks denote by $\{1, 2, \ldots, n\}$ and a continuum of risk-neutral outside financiers of unit mass. We index the representative outside financier — which may be another financial institution outside of the network of interest — by 0.

The economy lasts for three dates: $t = 0, 1, 2$. At the initial date, banks borrow funds from one another or the outside financiers to invest in projects that yield returns at the intermediate and final dates. More formally, each bank is endowed with k units of capital at $t = 0$ and has access to a project that requires an investment of size k to generate returns in future dates. However, in analogy to the "coconut" model of Diamond (1982), a bank cannot use its own funds to invest in its project, and instead, needs to borrow from other banks or the outside financiers. We assume that there are exogenous constraints on the extent to which banks can borrow from one another. Such restrictions may arise from liquidity and maturity mismatch across banks — for example, bank i may have excess funds to lend but not when bank j needs it — or may be due to asymmetric costs of peer monitoring, absence of long-term interbank relationships, or pairwise commitment problems.[11] Formally, we assume that

---

[10]For a recent survey of this literature, see Brunnermeier and Oehmke (2012).

[11]The assumption that bank i cannot invest its own funds in its project, and can only borrow from some specific banks, is meant to capture, in a simple way, the possibility that the investment opportunity and the funds required for undertaking that investment may not arise simultaneously. Consequently, banks may need to borrow when they have access to an investment opportunity, and can only do so from banks that have available funds exactly at the same time. As an alternative setup with identical implications, one can assume that date $t = 0$ is itself subdivided to multiple subperiods, in each of which some banks have excess capital while others have investment opportunities. In this alternative setup, the fact that i

bank j can borrow at most k_{ij} units of capital from bank i. This relationship may not necessarily be symmetric, in the sense that the extent to which bank j can borrow from bank i may be different from the extent to which the latter can borrow from the former. We define the lending/borrowing *opportunity network* as a weighted, directed graph on n vertices, where each vertex corresponds to a bank and a directed edge of weight k_{ij} from vertex j to vertex i is present if bank j has the opportunity to borrow from bank i. The existence of such a borrowing opportunity, however, does not necessarily imply that the banks would enter into a lending agreement. Rather, the interbank lending and borrowing decisions are endogenous. Banks may also decide to keep ("hoard") their funds, receiving a rate of return normalized to 1 (for example, by purchasing government bonds).

Finally, each bank can always borrow from the outside financiers, who are assumed to have sufficient funds at $t = 0$ with an opportunity cost of $r > 1$ between dates $t = 0$ and $t = 1$ (e.g., they have access to a linear risk-free technology with return r realized at $t = 1$).

Once bank i borrows k units of capital at $t = 0$, it invests in a project with a random short-term return of z_i at $t = 1$, and if held to maturity, a fixed, non-pledgeable long-term return of A at $t = 2$. The bank can liquidate the project prematurely at $t = 1$, at a loss. In particular, if the project is liquidated, the bank obtains a return of ζA, where $\zeta \approx 0$.[12] We assume that once a bank undertakes the project, it must also meet an outside obligation of magnitude $v > 0$ at $t = 1$, which is assumed to have seniority relative to the its obligations to other banks and the outside financiers. These more senior obligations may be senior debts (to other financial institutions outside the network), wages due to its workers or taxes due to the government.

2.2 Debt Contracts

Interbank lending takes place through debt contracts signed at $t = 0$.[13] We defer the complete description of the contracts and the process according to which agents enter into lending agreements to Section 4. For now, we take the amount of interbank borrowing and the interest rates as given. In particular, if ℓ_{ij} denotes the amount of capital borrowed by bank j from bank i, the face value of j's debt to i is equal to $y_{ij} = R_{ij}\ell_{ij}$, where R_{ij} is the corresponding interest rate. Note that by definition, $\ell_{ij} \leq k_{ij}$. The total debts (liabilities) of bank i at $t = 1$ is thus equal to $y_i + v$ where $y_i = \sum_{j \neq i} y_{ji}$.

Given our assumption that the long-term returns are not pledgeable, all debts have to be cleared at date $t = 1$. If bank j is unable to meet its $t = 1$ obligations in full, it defaults and has to liquidate its project prematurely where the proceeds are distributed among its creditors. We assume that all

---

cannot invest its funds in its own project, for example, can be interpreted as the assumption that the subperiods in which i has excess capital and the subperiods in which it has an investment opportunity do not coincide.

[12] This assumption ensures that liquidation of the project does not generate enough funds for the bank to meet its obligations.

[13] Even though we will allow these debt contracts to have covenants making interest rates contingent on the lending/risk-taking decisions of the borrowers, we do not allow general contracts. In a setting with bankruptcy and multiple creditors, general contracts pose certain conceptual problems which are beyond the scope of this paper (for example, obligations can be arranged so that they are highest when the bank is already bankrupt, thus making the creditor effectively senior relative to other creditors or so as to avoid the bankruptcy threshold in certain cases). In practice, the issue of counterparty risk is pertinent in the presence of debt contracts which are the norm in interbank markets.

junior creditors — that is, other banks and outside financiers — are of equal seniority. Hence, if bank j can meet its senior obligations, v, but defaults on its debt to the junior creditors, they are repaid in proportion to the face value of the contracts. On the other hand, if j cannot meet its more senior outside obligation v, its junior creditors receive nothing.

2.3 Financial Networks

The lending decisions of the banks and the resulting counterparty relations can be equivalently represented by an (endogenous) interbank network. In particular, we define the *financial network* corresponding to the bilateral debt contracts in the economy as a weighted, directed graph on n vertices, where each vertex corresponds to a bank and a directed edge from vertex j to vertex i is present if bank i is a creditor of bank j. The weight assigned to this edge is equal to y_{ij}, the face value of the contract between the two banks. By definition, the set of edges of a financial network is necessarily a subset of the set of edges of the underlying borrowing opportunity network.

A (financial) network is said to be *regular*, if all banks have identical interbank claims and liabilities; that is, $\sum_{j \neq i} y_{ij} = \sum_{j \neq i} y_{ji} = y$ for some y and all banks i. Even though the total liabilities and claims of all banks in a regular financial network are equal, the distribution of interbank claims may not be identical across different banks. Two important regular financial networks which feature prominently in our analysis are as follows.

Ring Financial Network   The *ring financial network*, also known as a *credit chain*, represents a configurations in which bank $i > 1$ is the sole creditor of bank $i - 1$ and bank 1 is the sole creditor of bank n; that is, $y_{i,i-1} = y_{1,n} = y$ for some y. Thus, as depicted in Figure 1(a), credit flows in only one direction. Though special, the ring financial network succinctly captures the properties of a financial network with very sparse connections.
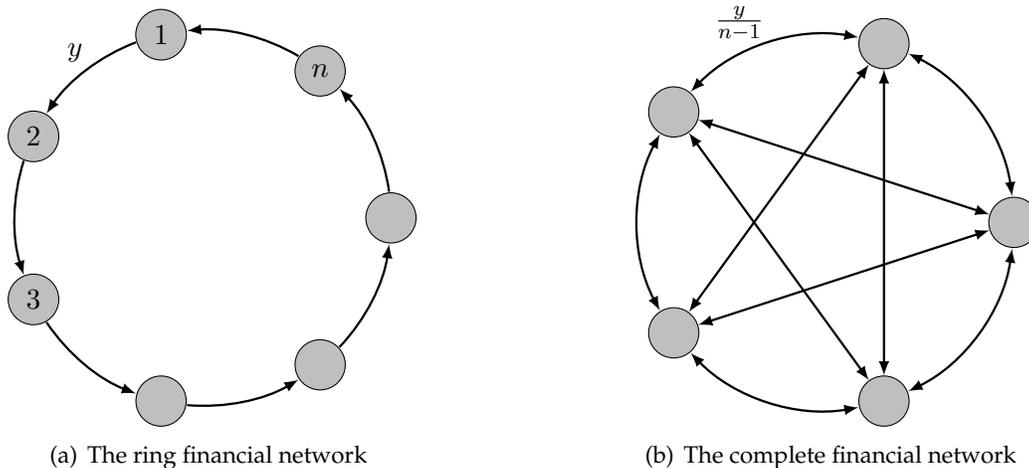

(a) The ring financial network

(b) The complete financial network

Figure 1: The ring and the complete financial networks

Complete Financial Network The polar opposite of the ring network is the financial network with fully diversified interbank lending, which we refer to as the *complete financial network*. In the complete financial network, depicted in Figure 1(b), all banks lend equally to all others; that is, $y_{ij} = y/(n-1)$ for all $i \neq j$. Given that the liabilities of each bank are spread across all other banks, the interbank connections in such an architecture are maximally "dense".

The ring and the complete financial networks correspond to the least and most densely interconnected financial networks, respectively. It is also useful to define a class of financial networks that exhibit intermediate degrees of interbank connectivity. We define a financial network to be the γ-*convex combination* of two regular financial networks corresponding to collections of pairwise contracts $\{y_{ij}\}$ and $\{\tilde{y}_{ij}\}$ if for all pairs of banks i and j, the face value of j's obligations to i is equal to $\gamma y_{ij} + (1 - \gamma)\tilde{y}_{ij}$. Thus, a financial network that is the γ-convex combination of the ring and the complete financial networks has an intermediate degree of density of connections. As γ decreases, such a network approaches the more densely connected complete network.

2.4 Payment Equilibrium

The ability of a bank to fulfill its promise to its creditors depends on the resources it has available to meet those obligations. In particular, the realized repayments by the bank on the debt to its creditors depends not only on the returns on its investment, but also on the realized value of repayments by the bank's debtors.

More formally, let x_{js} denote the repayment by bank s on its debt to bank j at $= 1$. By definition, $x_{js} \in [0, y_{js}]$. The total cash flow of bank j is then equal to $\alpha_j = c_j + z_j + \sum_{s \neq j} x_{js}$, where $c_j = k - \sum_{i \neq j} \ell_{ji}$ is the cash hoarded by the bank. If α_j is larger than the bank's total liabilities, $v + y_j$, then the bank is capable of meeting its obligations in full, and as a result, $x_{ij} = y_{ij}$ for all $i \neq j$. If, on the other hand, $\alpha_j < v + y_j$, bank j defaults and its creditors are repaid less than face value. In particular, when α_j is smaller than v, the bank defaults on its senior liabilities and its junior creditors receive nothing; that is, $x_{ij} = 0$. However, if $\alpha_j \in (v, v + y_j)$, the interbank payments by bank j would be proportional to the face value of the contracts. This is a consequence of the assumption that all junior creditors — which includes the creditor banks as well as the outside financiers — are of equal seniority and are repaid on *pro rata* basis. Thus, to summarize, the $t = 1$ payment of bank j to a creditor bank i is equal to

$$x_{ij} = \frac{y_{ij}}{y_j} \left[\min \left\{ y_j \, , \, e_j + \sum_{s \neq j} x_{js} \right\} \right]^+ , \tag{1}$$

where $e_j = c_j + z_j - v$ and $[\cdot]^+$ stands for $\max \{\cdot, 0\}$. Note that whenever the bank is unable to meet its obligations in full, it has to liquidate its project prematurely. The liquidation value, however, does not appear in (1) in light of the assumption that $\zeta \approx 0$.

Definition 1. Given cash holdings $\{c_j\}$, the face value of the bilateral interbank contracts $\{y_{ij}\}$, and the realizations of the shocks $\{z_j\}$, the interbank payments $\{x_{ij}\}$ form a *payment equilibrium* if they simultaneously solve (1) for all i and j.

A payment equilibrium is thus a collection of mutually consistent interbank payments at $t = 1$. The key observation is that a payment equilibrium captures the possibility of financial contagion in the system. In particular, given the interdependence of interbank payments across the financial network, a (sufficiently large) negative shock to a bank not only leads to that bank's default, but may also initiate a cascade of defaults by spreading to its creditors, its creditors' creditors, and so on. Our first result shows that the payment equilibrium is a well-defined notion.

Proposition 1. *For any given financial network and any realization of the shocks, a payment equilibrium always exists and is generically unique.*

Thus, for any given financial network, the payment equilibrium is uniquely determined over a generic set of parameter values and shock realizations.[14] The notion of payment equilibrium in our model is a generalization of the notion of a clearing vector introduced by Eisenberg and Noe (2001) and utilized by Shin (2008, 2009). Unlike the model of Eisenberg and Noe, the financial obligations of the banks in our model are of different seniorities.

Finally, for any given financial network and the corresponding payment equilibrium, we define the (utilitarian) social surplus in the economy as the sum of the returns to all agents; that is,

$$u = \pi_0 + \sum_{i=1}^{n} (\pi_i + T_i),$$

where $T_i \leq v$ is the transfer from bank i to its senior creditors, π_i is the bank's profit, and π_0 denotes the net return (in excess of their opportunity cost of r per unit of lending) to the outside financiers.

3 Financial Contagion

In this section, we study the repayment of interbank loans and the extent of financial contagion at $t = 1$, while taking the interbank debt contracts (signed at $t = 0$) as given. In particular, we focus on the properties of the payment equilibrium as a function of the structure of the financial network. We study the lending decisions of the banks and the formation of the financial network in Section 4.

In order to simplify the presentation, we restrict our attention to regular financial networks in which banks have no liabilities to outside financiers, and assume that all interbank loans are at the

---

[14] As we show in the proof of Proposition 1, in any connected financial network, the payment equilibrium is unique as long as $\sum_j (z_j + c_j) \neq nv$. In the non-generic case in which $\sum_j (z_j + c_j) = nv$, there may exist a continuum of payment equilibria, in almost all of which banks default due to "coordination failures". For example, if the economy consists of two banks with $c_1 = c_2 = v$, bilateral contracts of face values $y_{12} = y_{21}$ and no shocks, then defaults can occur if banks do not pay one another, even though both are solvent.

same interest rate, R.[15] Thus, the total interbank claims and liabilities of all banks are equal to $y = Rk$. Restricting our attention to regular financial networks enables us to focus on the relationship between the distributions of interbank liabilities and the extent of contagion, while abstracting away from effects that are driven by other features of the financial network, such as the asymmetry in the size of different banks' assets or liabilities.[16]

We further assume that the short-term returns on a given bank i's investment can only take two values $z_i \in \{a, a - \epsilon\}$, where $a > v$ is the return of the project in the "business as usual" regime and $\epsilon \in (a - v, a)$ corresponds to the magnitude of a negative shock to the project's returns. Finally, we assume that the realizations of the short-term returns are independent and identically distributed across different banks. These simplifying assumptions enable us to provide a meaningful comparison between the extent of financial contagion in different financial networks in a tractable manner. The following lemma characterizes the social surplus under the above assumptions.

Lemma 1. *Conditional on the realization of m negative shocks, the social surplus in the economy is equal to*

$$u = (n - \sharp defaults)A + na - m\epsilon.$$

Hence, the social surplus is simply determined by the number of defaults, which in turn reflects the extent of financial contagion. It is thus natural to measure the performance of a financial network in terms of the number of banks in default.

Definition 2. Conditional on the realization of m negative shocks,

  (i) the *stability* of a financial network is the inverse of the expected number of defaults.

 (ii) The *resilience* of the financial network is the inverse of the maximum number of possible defaults.

Thus, stability and resilience capture the expected and worst-case performances of the financial network in the presence of m negative shocks, respectively. Clearly, both measures of performance not only depend on the number (m) and the size (ϵ) of the realized shocks, but also on the structure of the financial network. To illustrate the relation between the extent of contagion and the financial network architecture in the most transparent manner, we initially assume that exactly one bank is hit with a negative shock. We generalize our results to the case of multiple shocks in Section 3.3.

3.1 Small Shock Regime

We first characterize the fragility of different financial networks when the size of the negative shock is relatively small.

---

[15]In the full equilibrium of the model described in Section 4, banks may face potentially different (and endogenously determined) interest rates depending on their probability of default.

[16]For example, Acemoglu et al. (2012) show that asymmetry in the degree of interconnectivity of different industries as input suppliers in the real economy plays a crucial role in the propagation of shocks.

Proposition 2. *Let $\epsilon^* = n(a - v)$ and suppose that $\epsilon < \epsilon^*$. Then, there exists y^* such that for $y > y^*$,*

(a) The ring network is the least resilient and least stable financial network.

(b) The complete network is the most resilient and most stable financial network.

(c) The γ-convex combination of the ring and complete networks becomes less stable and resilient as γ increases.

The above proposition thus establishes that as long as the size of the negative shock is below a critical threshold ϵ^*, the ring is the financial network most prone to financial contagion, whereas the complete network is the least fragile. Moreover, a more equal distribution of interbank obligations leads to less fragility.[17] Proposition 2 is thus in line with, and generalizes, the observations made by Allen and Gale (2000) and Freixas, Parigi, and Rochet (2000). The underlying intuition is that a more equal distribution of interbank liabilities implies that the burden of any potential losses is shared among more banks, creating a more robust system. In particular, in the extreme case of the complete financial network, the losses of a distressed bank are divided among as many creditors as possible, guaranteeing that the excess liquidity in the financial system can fully absorb the transmitted losses. On the other hand, in the ring financial network, the losses of the distressed bank — rather than being divided up between multiple counterparties — are fully transferred to its immediate creditor, leading to the creditor's possible default. A similar mechanism then guarantees that the distress is passed to a large fraction of the banks through a chain reaction, leading to a highly fragile system.

The condition that $\epsilon < \epsilon^* = n(a - v)$ requires the size of the negative shock to be less than the total "excess liquidity" available to the financial network as a whole. Recall that in the absence of any shock, $a - v$ is the liquidity available to each bank after meeting its senior obligations to outside the network. Proposition 2 also requires that interbank liabilities (and claims) are above a certain threshold y^*, which is natural given that for small values of y, no contagion would occur, regardless of the structure of financial network.

The extreme fragility of the ring financial network established by Proposition 2 is in contrast with the results of Acemoglu et al. (2010, 2012), who show that if the interactions over the network are linear (or log linear), the ring is as stable as any other regular network structure. This contrast reflects the fact that, with linear interactions, negative and positive shocks cancel each other out in exactly the same way independently of the structure of network. However, the often non-linear nature of financial interactions (captured in our model by the presence of debt contracts on which banks may default) implies that the effects of negative and positive shocks are not necessarily averaged out. Stability and resilience are thus achieved by minimizing the impact of the distress at any given bank on the rest of the system. The ring financial network is highly fragile precisely because the adverse ef-

---

[17]It may appear that part (c) of Proposition 2 can be proved by generalizing Lemma 6 of Eisenberg and Noe (2001). However, a closer inspection shows that the statement and the proof of the aforementioned lemma are incorrect. We provide a direct proof of Proposition 2 in the Appendix. We should stress again that neither Eisenberg and Noe (2001) nor any other paper we are aware of proves an equivalent statement.

fects of a negative shock to any bank are fully transmitted to the bank's immediate creditor, triggering maximal financial contagion.

3.2 Large Shock Regime

Proposition 2 shows that as long as the magnitude of the negative shock is below the threshold ϵ^*, a more equal distribution of interbank liabilities leads to less fragility. In particular, it shows that the complete network is the most stable and resilient financial network: except for the bank that is directly hit with the negative shock, no other bank defaults. Our next result, however, shows that when the magnitude of the shock is above the critical threshold ϵ^*, this picture changes dramatically.

A collection of banks $M \subset N$ is said to form a δ-*component* of the financial network if (i) the total obligations of banks outside of M to any bank in M is at most $\delta \geq 0$; and (ii) the total obligations of banks in M to any bank outside of M is no more than δ. Intuitively, for small values of δ, banks in a δ-component have weak ties to the rest of the financial network. We say a financial network is δ-*connected* if it contains a δ-component.

Proposition 3. *Suppose $\epsilon > \epsilon^*$ and $y > y^*$. Then,*

(a) The complete and the ring networks are the least stable and least resilient financial networks.

(b) For small enough values of δ, any δ-connected financial network is strictly more stable and resilient than the ring and complete financial networks.

Thus, when the magnitude of the negative shock is sufficiently large, the complete network exhibits a form of *phase transition*: it flips from being the most to the least stable and resilient network, achieving the same level of fragility as the ring network. In particular, when $\epsilon > \epsilon^*$, all banks in the complete network default. The intuition behind this result is simple: given that all banks in the complete network are creditors of the distressed bank, the adverse effects of the negative shock are *directly* transmitted to them. Thus, when the size of the negative shock is large enough, not even the originally non-distressed banks are capable of paying their debts in full, leading to the default of all banks.

Not all financial systems are as fragile in the face of large shocks. Instead, as part (b) shows, if the financial network contains a δ-component for small enough values of δ, then it is strictly more stable and resilient than both the complete and the ring networks. The presence of such "weakly connected" components in the network guarantees that the losses — rather than being transmitted to all other banks — are borne in part by the distressed banks' senior creditors.

Taken together, Propositions 2 and 3 illustrate the "robust-yet-fragile" property of highly interconnected financial networks conjectured by Haldane (2009). They show that more densely interconnected financial networks, epitomized by the complete network, are more stable and resilient in
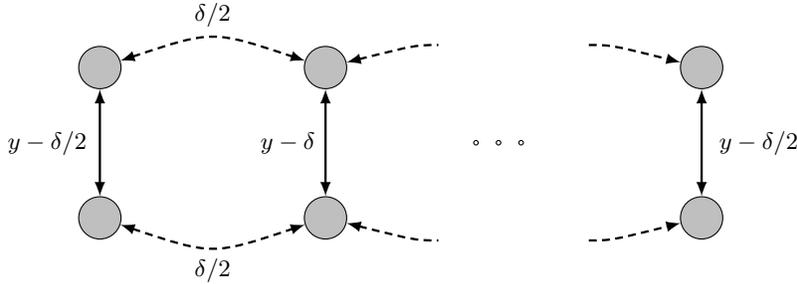
Figure 2: A collection of δ-connected pairs of banks

response to a range of shocks. However, once we move outside this range, these dense interconnections act as a channel through which shocks to a subset of the financial institutions transmit to the entire system, creating a vehicle for instability and systemic risk.

The intuition behind such a phase transition it related to the presence of two types of "shock absorbers" in our model, each of which capable of reducing the extent of contagion in the network. The first absorber is the excess liquidity $a - v > 0$ of non-distressed banks at $t = 1$: the impact of a shock is attenuated once it reaches banks with excess liquidity. This mechanism is utilized more effectively when the financial network is more "complete", an observation in line with the results of Allen and Gale (2000) and Freixas, Parigi, and Rochet (2000). However, the claim v of senior creditors of the distressed banks also function as a shock absorption mechanism. Rather than transmitting the shocks to other banks in the system, the senior creditors can be forced to bear (some of) the losses and hence, limit the extent of contagion. In contrast to the first mechanism, this shock absorber is best utilized in weakly connected configurations and is the least effective in the complete financial network. Thus, when the shock is so large that it cannot be fully absorbed by the excess liquidity in the system — which is exactly when $\epsilon > \epsilon^*$ — financial networks that significantly utilize the second absorber are less fragile.

Example 1. Consider the financial network depicted in Figure 2, consisting of $n/2$-many δ-components of size two, where $\delta < a - v$. This configuration is reminiscent of the traditional "unit banking system" with minimal connections between banks across different regions or industries.

If $2(a - v) < \epsilon < \epsilon^*$, a negative shock to a bank would guarantee the default of its (sole) major counterparty. Thus, in the face of a small shock, the financial network is *strictly* less stable and resilient than the complete financial network, in which only one bank defaults. On the other hand, given the absence of any significant inter-pair claims and liabilities, a negative shock to a bank would not cause any other defaults besides that of its major counterparty, regardless of the size of the shock. Hence, if $\epsilon > \epsilon^*$, it is strictly more stable and resilient than the complete network. In fact, in the large shock regime, the financial network in Figure 2 is the most stable and resilient financial network.

3.3 Multiple Shocks

The insights on the relationship between the extent of contagion and the structure of the financial network studied so far generalize to the case of multiple negative shocks.

Proposition 4. *Let m denote the number of negative shocks and let $\epsilon_m^* = n(a - v)/m$. There exist constants $y_m^* > \hat{y}_m > 0$, such that*

(a) *If $\epsilon < \epsilon_m^*$ and $y > y_m^*$, then the complete network is the most stable and resilient financial network, whereas the ring network is the least resilient.*

(b) *If $\epsilon > \epsilon_m^*$ and $y > y_m^*$, then the complete and the ring financial networks are the least stable and resilient financial networks. Furthermore, for small enough values of δ, any δ-connected financial network is strictly more stable than the complete and ring financial networks.*

(c) *If $\epsilon > \epsilon_m^*$ and $y \in (\hat{y}_m, y_m^*)$, then the complete financial network is the least stable and resilient financial network. Furthermore, the ring financial network is strictly more stable than the complete financial network.*

Parts (a) and (b) generalize the insights of Propositions 2 and 3 to the case of multiple negative shocks. The key new observation is that the critical threshold ϵ_m^* that defines the boundary of the small and large shock regimes is a decreasing function of m. Thus, the number of negative shocks plays a role similar to that of the size of the shocks. More specifically, as long as the magnitude *and* the number of negative shocks affecting financial institutions are sufficiently small, more complete interbank claims enhance the stability and the resilience of the financial system. This is due to the fact that the more complete the interbank connections are, the better the excess liquidity of non-distressed banks are utilized in absorbing the shocks. On the other hand, if the magnitude *or* the number of shocks are large enough so that the excess liquidity in the financial system is not sufficient for absorbing the losses, financial interconnections serve as a propagation mechanism, creating a more fragile financial system. Weakly connected networks ensure that the losses are shared with the senior creditors of the distressed banks, protecting the rest of the system.

Part (c) of Proposition 4 contains another new result. It shows that in the presence of multiple shocks, the claims of the senior creditors in the ring financial network are used more effectively as a shock absorption mechanism than in the ring financial network. In particular, the closer the banks hit with the negative shocks are to one another in the ring financial network, the larger the loss their senior creditors are collectively forced to bear. This limits the extent of contagion in the network.

Finally, we remark that even though we illustrated our results by focusing on an environment in which shocks can take only two values, similar results can be obtained for more general shock distributions.

3.4 Illustrative Simulations

We end this section by providing a few simulations, illustrating the performance of different financial network structures as a function of the size of the negative shock.

We focus on four different financial network configurations consisting of $n = 20$ banks: (i) the complete financial network; (ii) the ring financial network; (iii) the γ-convex combination of the ring and the financial networks for $\gamma = 3/4$; and (iv) the *double-clique* financial network which is a weakly connected network consisting of two 0-components (that is, disconnected components) of size 10, each of which is complete. We vary the size of the negative shock ϵ and simulate the expected number of defaults in each of the above-mentioned financial networks. For the purpose of the simulations, we let $y = 30$ and $a - v = 1$, which implies that the critical threshold is $\epsilon^* = 20$.

Figure 3(a) depicts the expected number of defaults conditional on the realization of a single negative shock of size ϵ. As it is evident form the graph, the ring financial network is the most fragile regardless of the size of the shock. The figure also clearly highlights the phase transition of the complete financial network: once the size of the shock crosses ϵ^*, the complete financial network flips from the most to the least stable architecture. The financial network that is a γ-convex combination of the ring and complete financial networks exhibits an intermediate level of robustness. Finally, the results for the double-clique financial network highlights the fact that weak connections limit the extent of financial contagion in the face of large shocks. These observations are in line with Propositions 2 and 3. As a final remark note that the expected number of defaults in the double-clique financial network exhibits a discontinuity at $\epsilon = 10$. Recall that this financial network consists of two complete cliques of size 10. Thus, in line with the results for the complete network, each of these cliques would go through a phase transition as the size of the shock crosses $\hat{\epsilon} = 10(a - v) = 10$.

(a) Single shock $m = 1$ (b) Independent shocks with $p = 0.05$
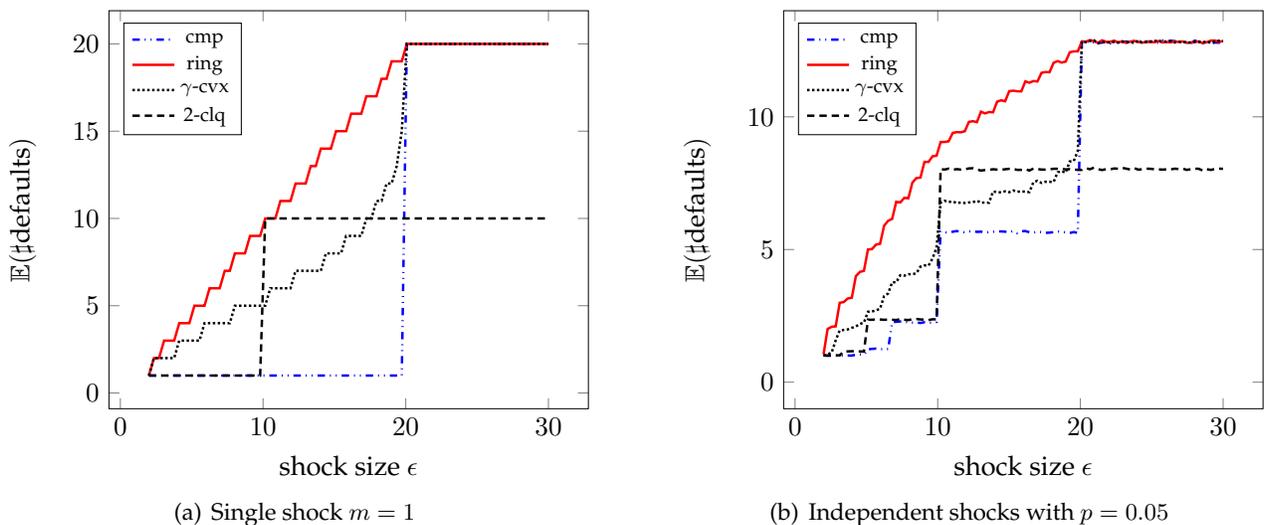
Figure 3: Number of failures as a function of the size of the shock

Figure 3(b) depicts the extent of financial contagion when the shock realizations are independent and identically distributed, with each bank being hit with a negative shock of size ϵ independently with probability $p = 0.05$. This figure confirms that our results do not depend on a specific number of shocks (at random) hitting one of the banks; once again the ring financial network is the most fragile of the four networks, the complete financial network outperforms all others for $\epsilon < \epsilon^*$, and the weakly connected double-clique financial network is the most stable for $\epsilon > \epsilon^*$.

4   Financial Network Externality

In the previous section, we took the lending decisions of the banks as given and studied the extent of financial contagion as a function of the structure of the financial network. In this section, we endogenize the interbank counterparty relations and study financial networks that arise in equilibrium. The key endogeneity involves the structure and the terms of bilateral interbank agreements. Our main results are to establish the presence of a financial network externality and study the inefficiencies that it introduces.

We begin by describing how the terms of trade across banks are determined. At the beginning of date $t = 0$, all banks and outside financier simultaneously post contracts detailing the terms at which they are willing to lend to one another. We assume that the posted contracts are *standard debt contracts with contingency covenants*, according to which the lender specifies the interest rates at which it lends to the borrower as a function of the borrower's lending behavior. Given the posted contracts, each bank then decides how much and from whom to borrow, which in turn determines the interbank interest rates. More formally, the timing of events over date $t = 0$ is as follows.

1. All agents $i \in \{0, \ldots, n\}$ simultaneously post contracts of the form $\mathbf{R}_i = (R_{i1}, \ldots, R_{in})$, where R_{ij} is a mapping from the bank j's lending decision $(\ell_{j1}, \ldots, \ell_{jn})$ to the interest rate on bank j's debt to i. If i cannot lend to bank j or decides not to do so, then $R_{ij} = \varnothing$.

2. After observing the set of posted contracts, each agent can withdraw one or more of its contract offers if it so wishes. The final set of contracts offered by bank i is thus $\hat{\mathbf{R}}_i = (\hat{R}_{i1}, \ldots, \hat{R}_{in})$, where $\hat{R}_{ij} \in \{R_{ij}, \varnothing\}$.[18]

3. Given the set of contracts $(\hat{\mathbf{R}}_0, \ldots, \hat{\mathbf{R}}_n)$, each bank j decides on the amount b_{ij} that it borrows from agent i.

By construction, the borrowing decisions of the banks determine the amount of interbank lending; that is, $\ell_{ij} = b_{ij}$. This, in turn, pins down the interest rate on bank j's debt to i at $\hat{R}_{ij}(\ell_{j1}, \ldots, \ell_{jn})$.

---

[18]The second stage of the game, in which contract offers can be withdrawn, is introduced in order to rule out certain unnatural equilibria that may arise due to "coordination failures". In particular, unless bank i can withdraw its posted contracts, it cannot make its lending decisions contingent on the contract posted by a potential creditor bank s. In other words, without the possibility of contract withdrawals, once bank i posts a contract R_{ij}, it already commits to lend to bank j regardless of the value of R_{si}.

Given the large set of possible lending decisions by the banks, the contract $\mathbf{R}_i$ posted by bank i is an infinite dimensional object. In order to simplify the exposition and derivation of our main results, unless otherwise noted, we restrict the set of interbank borrowings by assuming that $b_{ij} \in \{0, k_{ij}\}$. That is, if banks i and j enter into a lending agreement, then the borrowing would be equal to the maximal borrowing capacity. This simplification reduces each borrowing decision to a binary choice. Consequently, the contract $R_{ij}(\ell_{j1}, \ldots, \ell_{jn})$ is reduced to a vector, specifying the interest rates at which bank i is willing to lend to j as a function of the identities of j's counterparties. With some abuse of notation, we use R_{ij} to denote both the contingent debt contract between i and j, as well as the actual interest rate on j's debt to i once all the bilateral interbank agreements are finalized. Also, whenever there is no risk of confusion, we use R_{ij} and $\hat{R}_{ij}$ interchangeably.

Definition 3. A *full (subgame perfect) equilibrium* is a collection of contracts posted by the banks and the outside financiers, given by $(\mathbf{R}_0, \mathbf{R}_1, \ldots, \mathbf{R}_n)$ and $(\hat{\mathbf{R}}_0, \hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_n)$, bilateral borrowing decisions $\{b_{ij}\}$, and interbank repayments $\{x_{ij}\}$ such that,

(a) Given the financial network, the repayments on the loans are determined by the corresponding payment equilibrium.

(b) Given the posted contracts, the financial network is a Nash equilibrium of the corresponding subgame.[19]

(c) Neither the banks nor the outside financiers have an incentive to deviate by withdrawing or posting a different contract at any stage of the game at time $t = 0$.

A financial network $\{y_{ij}\}$ is thus part of an equilibrium if (i) taking the interest rates as given, the banks have no incentive to unilaterally change their counterparties; and (ii) they cannot make strictly higher profits by charging different interest rates. Given that the outside financiers are risk neutral and act competitively, the equilibrium contract $\mathbf{R}_0$ gives them an expected return equal to their opportunity cost, r.

The most important feature of our setup is that equilibrium interest rates are determined endogenously. In particular, given that the lending behavior of a bank may expose its creditors to additional counterparty risks, the presence of covenants that make interest rates contingents on the borrower's behavior forces the banks to internalize the impact of their decisions on their immediate creditors. This feature, as we show in the next subsection, ensures that the most obvious form of bilateral externalities are internalized, thus providing a useful framework for analyzing financial network externalities. Throughout, we use a notion of efficiency according to which the social planner controls

---

[19] An alternative is to use a solution concept similar to the pairwise stability notion of Jackson and Wolinsky (1996), often used in problems of network formation. Given the more specific context here, posting of interest rates combined with subgame perfection provides a powerful solution concept that is both more transparent and easier to work with. Our main results and the insights that follow are robust with respect to the choice of the solution concept.

the lending decisions, but not the interbank interest rates, which are determined by the equilibrium behavior of the outside financiers.[20]

4.1 Bilateral Efficiency: The Three-Chain Financial Network

To show how bilateral financial externalities are internalized thanks to the covenants, we focus on a special architecture in which no other form of network externalities is present. In particular, we consider an economy comprising of three banks labeled $\{1, 2, 3\}$, each endowed with k units of capital. In order to invest in their projects, banks 1 and 2 need to borrow k units of capital from banks 2 and 3, respectively. That is, $k_{21} = k_{32} = k$, and $k_{ij} = 0$ otherwise. Bank 3, on the other hand, does not borrow and simply acts as a (potential) lender to bank 2. The assumption that there are no other banks or financiers lending to bank 3 rules out the possibility of network externalities.[21]

Given the above description, if no bank relies on the outside financiers for funding, the three-chain financial network depicted in Figure 4 would form. To further simplify exposition, we assume that bank 1 is the only bank subject to a negative shock. In particular, we assume that $z_1 \in \{a - \epsilon, a\}$ where the negative shock is realized with probability p and satisfies $2(a - v) < \epsilon < 2(a - v) + k$. Banks 2 and 3, on the other hand, are not subject to shocks, i.e., $z_2 = z_3 = a$ with probability one.
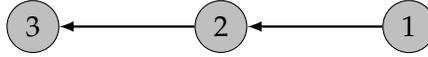

Figure 4: The three-chain financial network.

Given that the returns on bank 1's investments are subject to negative shocks, bank 3's profits depend on whether bank 2 decides to lend to bank 1. In particular, a bilateral lending agreement between banks 1 and 2 not only increases the default probability of bank 2, but also exposes bank 3 to the risk of contagion. Yet, our next result shows that thanks to the contingency covenants in the debt contracts, bank 2 fully internalizes the effect of its lending decisions on bank 3.

Proposition 5. *The three-chain financial network is an equilibrium if and only if it is efficient.*

The above proposition thus establishes that each bank takes the effects of its actions on its immediate creditors into account. In other words, *bilateral* externalities are internalized because a bank's creditors can offer contracts whose terms induce the "right" behavior for the borrower. In the particular case of the three-chain financial network, if a bilateral lending agreement between banks 1 and 2 increases the risk of contagion to bank 3 beyond the socially efficient level, bank 3 would be willing

---

[20]If we allow the social planner to also determine the interbank interest rates, she can increase utilitarian social welfare by setting all interest rates equal to zero, thus minimizing defaults. Such a framework, clearly, does not constitute a reasonable benchmark for comparison.

[21]This assumption implicitly implies that bank 3 is not subject to default. This has no bearing for the results that follow. To restore symmetry between the three banks, we can alternatively assume that bank 3 also has access to a project it can invest in and loses an extra amount of A (due to costly liquidations) if its returns fall below a certain threshold. Proposition 5 remains valid under this assumption.

to offer bank 2 a sufficiently low interest rate provided that the latter hoards cash, hence, ensuring efficiency.

4.2 Financial Network Externality: Overlending

We next show that in the presence of counterparty risk and the possibility of financial contagion, private and public incentives for forming financial connections do not generally coincide. In particular, we start by showing that the equilibrium often features "overlending", in the sense that banks lend to one another even though the social planner would have preferred that they hoarded cash.

To illustrate the nature of these inefficiencies in a clear fashion, we focus on a special configuration of interbank lending opportunities. We assume that bank $i \in \{1, \ldots, n-1\}$ cannot borrow from any bank other than bank $i + 1$, whereas bank n can only borrow from bank 1; that is, $k_{1,n} = k_{i,i-1} = k$ and $k_{ij} = 0$ otherwise. Thus, if all projects are fully financed via interbank loans, the ring financial network, depicted in Figure 1(a) would form. We assume that a single shock of size ϵ hits the banks uniformly at random.

Proposition 6. *Suppose that $\epsilon < \epsilon^*$. Then, there exist constants $\underline{\alpha} < \bar{\alpha}$ such that,*

(a) The ring financial network is part of an equilibrium if $\underline{\alpha}A < (r-1)k$.

(b) The ring financial network is socially inefficient if $(r-1)k < \bar{\alpha}A$.

The ring financial network is thus part of an equilibrium if the interest rates that the banks can charge when they lend to one another — the benefit of which over hoarding cash is $(r-1)k$ — is large enough to justify the subsequent increase in the expected cost of a default, which is proportional to A. On the other hand, the ring financial network would be socially inefficient whenever the costs associated with the higher risk of financial contagion are high enough so that it is no longer justified for all banks to lend. The more important consequence of Proposition 6 is that, unlike the case of the three-chain financial network, the equilibrium and efficiency conditions no longer coincide: as long as $\underline{\alpha}A < (r-1)k < \bar{\alpha}A$, the ring financial network is formed in equilibrium, even though it is inefficiently fragile.

The juxtaposition of Propositions 5 and 6 clarifies that the inefficiency of equilibrium financial networks is not due to a simple bilateral externality. Rather, it is due to the presence of a *financial network externality*: lending by each bank creates a pathway for the translation of idiosyncratic shocks into financial contagion and systemic crises . Even though, thanks to the covenants, each bank takes the effect of its actions on itself and its immediate creditors into account, it does not fully internalize the effects of its decisions on the its creditors' creditors and so on. For example, in the case of the ring financial network, despite the fact that the interest rate $R_{i+1,i}$ faced by bank i depends on whether it decides to lend or not, the effects of i's actions on banks other than $i + 1$ are not reflected in that interest rate. In particular, neither i nor $i + 1$ take into account that lending by bank i may lead to a cascade of defaults of length $\tau > 2$.

The presence of this type of financial network externality implies that financial stability is a public good, which cannot be resolved by bilateral contracting and is only inadequately provided in equilibrium.[22]

4.3   Financial Network Externality: Insufficiently Dense Networks

We next show that the financial externality identified in the previous subsection may lead to the formation of insufficiently dense networks, in the sense that banks may lend to the "wrong" set of banks from the social planner's point of view.

To illustrate this possibility, we focus on an economy in which each bank can lend to two different borrowers. In particular, consider an n-bank economy (where n is even) in which banks labeled $2i$ and $2i-1$ can lend to banks labeled $2i-2$ and $2i-3$; that is, $k_{2i,2i-2} = k_{2i,2i-3} = k_{2i-1,2i-2} = k_{2i-1,2i-3} = k$ and $k_{ij} = 0$ otherwise. Thus, if all banks decide to lend equally to their potential borrowers, the *interlinked rings* financial network, depicted in Figure 5(b), would form. However, each bank can also decide to follow an undiversified lending strategy and instead lend to only one borrower. Following such a strategy by all banks would lead to the formation of the *double-ring* financial network depicted in Figure 5(a). We assume that a single shock of size ϵ hits the banks uniformly at random.

Proposition 7. *Suppose that $\epsilon < \epsilon^*/2$ and let $\tau = \lceil \epsilon/(a-v) \rceil - 1$ be the number of defaults in the double-ring financial network. Then, there exists $\underline{\alpha} > 0$ such that*

(a) The double-ring financial network is part of an equilibrium if $\underline{\alpha}A < (r-1)k$.

(b) The double-ring financial network is socially inefficient if τ is even.

Thus, for large enough values of r, the double-ring financial network is part of an equilibrium even though it is not socially efficient. As in Proposition 6, the assumption that $\underline{\alpha}A < (r-1)k$ guarantees that no bank in the double-ring financial network has an incentive to deviate by hoarding cash. Unlike the previous example, each bank can also deviate by following a diversified lending strategy. Yet, as part (a) shows, such a deviation would not be profitable either: for a large enough value of τ, lending to a diversified set of banks would expose the bank (and its immediate creditor) to a higher level of counterparty risk.

The key observation is that even though no bank finds diversified lending profitable, such a deviation imposes a positive externality on the rest of the system. Recall that by Proposition 2, the extent of financial contagion is reduced the more densely interconnected the financial network is. Hence, following a diversified lending strategy would benefit other banks as the excess liquidity in the system is more efficiently utilized in absorbing the shocks. This is the source of inefficiency of the double-ring financial network established in part (b) of Proposition 7.[23]

---

[22]As noted in the Introduction, this result is related to Zawadowski (2012) who shows that banks may underinvest in default insurance on their counterparties, partly because they do not internalize the positive spillovers that such insurance would create to other banks. Though clearly complementary, the exact mechanism and theoretical analysis in his paper are
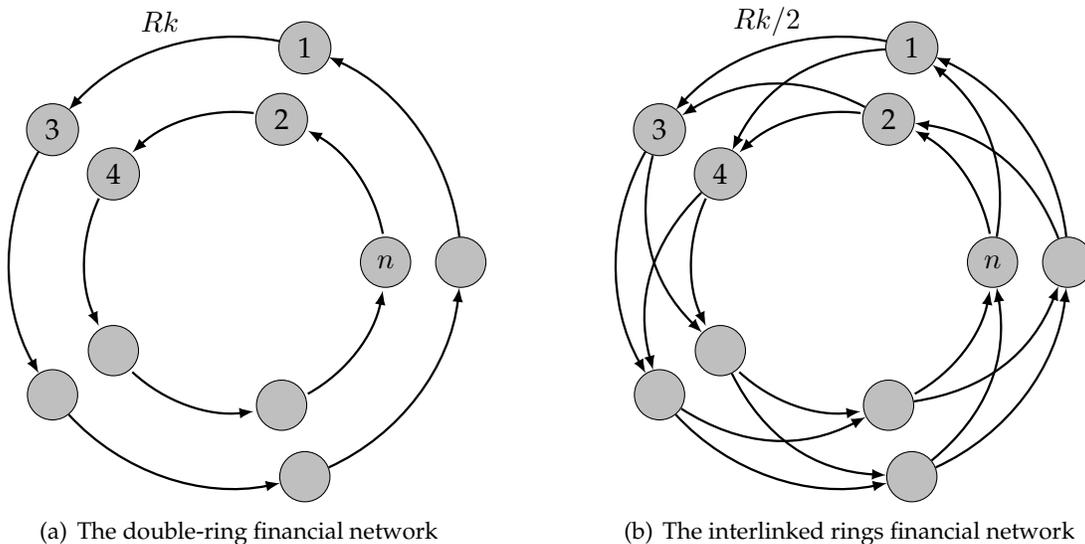
(a) The double-ring financial network

(b) The interlinked rings financial network

Figure 5: The double-ring and the interlinked rings financial networks

4.4 Robust-Yet-Fragile Financial Networks

We end this section by showing that the the same type of financial externalities can also lead to the emergence of robust-yet-fragile financial networks, in the sense that the entire financial network remains (inefficiently) vulnerable to the realization of a rare negative shock.

To this end, we focus on the particular configuration of interbank lending opportunities described by $k_{ij} = k/(n-1)$ for all $i \neq j$; that is, if no bank hoards cash, the complete financial network depicted in Figure 1(b) would form. We once again assume that a single shock hits the banks uniformly at random. The shock can take two distinct values ϵ_h and ϵ_ℓ with probabilities p and $1 - p$, respectively, where $\epsilon_\ell < \epsilon^* < \epsilon_h$. Thus, for small enough values of p, the two realizations correspond to "large but rare" and "small but frequent" negative shocks, respectively. Our next result provides necessary and sufficient conditions for the complete financial network to be socially efficient and/or part of an equilibrium.

Proposition 8. *Suppose that $\epsilon_\ell < \epsilon^* < \epsilon_h$. There exist constants $\bar{p} > 0$ and $\xi > 0$ such that*

(a) If $p = 0$, the complete financial network is socially efficient.

(b) If $p > 0$ and $pA > \xi$, the complete financial network is socially inefficient.

(c) If $p < \bar{p}$, the complete financial network is part of an equilibrium.

The intuition behind part (a) is simple. Recall from Proposition 2 that in the presence of only small shocks, the complete financial network is the most stable financial network and ensures that only the

---

quite different from ours.

[23]The requirement that τ is an even number is a simple technical assumption guaranteeing that the extent of financial contagion in the interlinked rings financial networks is strictly smaller than τ.

21

bank that is hit with the negative shock defaults. This also implies that when $p = 0$, the social surplus is maximized whenever all banks fully lend to one another, forming the complete network. This is no longer true, however, when there is a positive probability of a large shock realization. As part (b) of Proposition 8 shows, even for small (yet positive) values of p, the complete financial network is socially inefficient if A is large enough. The reason behind this inefficiency can be understood in light of Proposition 3. Recall that conditional on the realization of a large shock, all banks in the complete financial network default. Thus, if the cost of a default is sufficiently large, the social planner would prefer a weakly connected network or the empty configuration — in which banks withhold lending altogether — to the complete financial network. Such a weakly connected or fully disconnected (empty) architecture would ensure that the realization of a large shock at some distant corner of the network does not translate into a systemic crises, affecting all banks. Finally, part (c) of Proposition 8 shows that, if p is small enough, the complete financial network is part of an equilibrium. This is due to the fact that if the large shock, and hence, financial contagion is fairly unlikely, the banks find lending more profitable than hoarding cash.

The key insight in Proposition 8 is obtained from comparison of the conditions in the three parts. Taken together, parts (a) and (c) imply that the complete financial network is part of an efficient equilibrium whenever the possibility of a large systemic shock is ruled out (because when $p = 0$, the complete financial network fully absorbs the shock without any contagion). Public and private incentives for interbank lending diverge, however, as soon as a large shock is possible considered impossible. This can be seen by comparing statements (b) and (c). As far as an individual bank is concerned, hoarding cash (as opposed to lending to the other banks) decreases its default probability by $p(n-1)/n$. But, hoarding cash by some bank i also protects the rest of banking system whenever i is hit with the large shock. In fact, from the social planner's point of view, hoarding cash by a bank decreases the expected number of defaults by $2p(n-1)/n$. The presence of the financial network externality thus implies that the possibility of a rare, large shock may lead to the formation of (inefficiently) robust-yet-fragile financial networks *a la* Haldane (2009).[24]

5   Renegotiations and Financial Contagion

In this section, we relax the assumption that long-term returns are non-pledgeable, thus allowing banks to pledge at least a fraction of their long-term (date $t = 2$) returns to their creditors. In particular, we assume that a fraction $\kappa \in [0, 1]$ of the long-term returns can be pledged at $t = 1$, essentially allowing banks to pass the ownership of a fraction κ of their long-term returns to their creditors.[25] Because these returns remain non-pledgeable at $t = 0$, banks still need to fund their projects through

---

[24]The comparison between statements (b) and (c) of Proposition 8 implies the emergence of the complete financial network as an inefficient equilibrium only if the intersection of the conditions in the two statements is non-empty; that is, only if $p < \bar{p}$ and $pA > \xi$ can hold simultaneously. It can be verified that for large enough n, the value of $\bar{p}$ becomes independent of A. Hence, for any $p \in (0, \bar{p})$, there exists a large enough A such that $pA > \xi$ holds.

[25]We are thus assuming that two banks cannot renegotiate with one another, unless they are counterparties. We also continue to assume that the banks cannot pledge their long-term returns to the outside financiers.

standard debt contracts that clear at the intermediate date. The difference, however, arises in the event that a bank is unable to make its debt payments at $t = 1$, in which case it can pledge up to κA of its long-term returns.[26]

In order to formalize the renegotiation stage, we assume that the intermediate date is subdivided into M periods (where $M > 2n^2$) during which banks make renegotiation offers to their counter-parties. In particular, first bank 1 makes renegotiation offers to its creditor banks in a pre-specified order, followed by renegotiation offers by bank 2 to its own creditors, and so on.[27] These renegotiation offers take the form of a bank i pledging some amount $A_{ji} \geq 0$ to its creditor j, subject to the constraint that its pledges do not exceed κA. In exchange, i may request postponement of its debt repayments or even injection of additional funding from j. The interbank payments are made at the end of date $t = 1$ after the renegotiation phase is over. For a given regular financial network, we have the following result:

Proposition 9. *Suppose that the financial network is connected.*

(a) If $\epsilon < \min\{\epsilon_m^*, \kappa A/m + a - v\}$, then no bank defaults.

(b) If $\epsilon > \min\{\epsilon_m^*, \kappa A + a - v\}$, then at least one bank defaults.

Thus, for large enough values of κ, the possibility of renegotiations by counterparties would prevent all defaults, even if some banks are distressed. On the other hand, if κ is small enough, then distressed banks cannot avoid costly defaults by renegotiating their obligations to counterparties in exchange for future payments. In fact, no renegotiations occur throughout the network, all distressed banks default and financial contagion occurs as if $\kappa = 0$. These observations are intuitive. More interestingly, however, is the case that a single bank is hit with a negative shock. In particular, if $m = 1$, Proposition 9 is essentially another phase transition result: there exists a threshold

$$\hat{\epsilon} = \min\{\epsilon^*, \kappa A + a - v\},$$

such that for $\epsilon < \hat{\epsilon}$, the structure of the financial network (beyond connectivity) plays no role whatsoever in the extent of contagion. In fact, regardless of the detailed structure of the financial network, the excess liquidity of the banking system — through a sequence of renegotiations — is utilized to avoid all bankruptcies. This picture is dramatically different if ϵ is above the critical threshold $\hat{\epsilon}$, in which case even all the liquidity that can be mobilized is insufficient for preventing defaults. Thus, failures would occur and local network effects similar to those characterized in the Section 3 would become the defining factor in the extent of financial contagion.

Note that for the above phase transition result to hold we do not need to assume that a bank can renegotiate with all other banks. In fact, as long as each bank can renegotiate with its creditors (and

---

[26]Equivalent results would be obtained if the initial loans are made with more complex contracts that make payments both at date $t = 1$ and $t = 2$. Sticking to the standard debt contracts and allowing renegotiation is just a convenient convention.

[27]The same results would obtain with simultaneous renegotiation offers if we pick the best Nash equilibrium of the simultaneous offer game.

the financial network is connected), the excess liquidity at any part of the banking system can be diverted to the distressed bank. This is due to the fact that even if a distressed bank's creditors do not have enough funds to lend, they can themselves pledge their own long-term returns to their creditors, and raise the necessary funds to lend to the distressed bank. Such renegotiation chains guarantee that excess liquidity would be allocated efficiently.

6   Conclusions

The recent financial crisis has rekindled interest in the relationship between the structure of the financial network and systemic risk. Two polar views on this relationship have been suggested in the academic literature and the policy world. The first maintains that the "incompleteness" of the financial network can be a source of instability, as individual banks are overly exposed to the liabilities of a handful of financial institutions. Thus, according to this argument, a more complete financial network which limits the exposure of the banks to any one counterparty would be less prone to systemic failures. The second view, in stark contrast, hypothesizes that it is the highly interconnected nature of the financial system that contributes to its fragility, as it facilitates the spread of financial distress and solvency problems from one bank to the rest in an epidemic-like fashion.

This paper provides a tractable theoretical framework for the study of the economic forces shaping the relationship between the structure of the financial network and systemic risk. We show that as long as the magnitude (or the number) of negative shocks is below a critical threshold, a more equal distribution of interbank obligations leads to less fragility. In particular, all else equal, the sparsely connected ring financial network (corresponding to a credit chain) is the most fragile of all configurations, whereas the highly interconnected complete financial network is the configuration least prone to contagion. In line with the observations made by Allen and Gale (2000), our results establish that, in the more complete networks, the losses of a distressed bank are passed to a larger number of counterparties, guaranteeing a more efficient use of the excess liquidity in the system in forestalling defaults.

We also show that when negative shocks are larger than a certain threshold, the second view on the relationship between the structure of the financial network and the extent of contagion prevails. In particular, completeness is no longer a guarantee for stability. Rather, in the face of large shocks, financial networks in which banks are only weakly connected to one another would be less prone to systemic failures. Such a "phase transition" is due to the fact that, the senior liabilities of banks, as well as the excess liquidity within the financial network, can act as shock absorbers. Weak interconnections guarantee that the more senior creditors of a distressed bank bear most of the losses and hence, protect the rest of the system against cascading defaults. Our model thus formalizes the robust-yet-fragile property of interconnected financial networks conjectured by Haldane (2009). On a broader level, our results highlight the possibility that the same features that make a financial network structure more stable under certain conditions may function as significant sources of systemic risk and instability

under another.

The paper also provides an analysis of the efficiency of the financial networks that emerge in equilibrium. We assume that interbank interest rates adjust endogenously in order to reflect the risk that each bank's decisions impose on its counterparties. Nevertheless, we show that equilibrium financial networks may be excessively prone to the risk of financial contagion. The intuition underlying this result is that even though banks fully internalize the effects of their lending (and risk-taking) decisions on their immediate creditors, they do not take into account the fact that their lending decisions may also put many other banks (such as their creditors' creditors) at a greater risk of default. Thus, our results highlight the presence of a *financial network externality*, which cannot be internalized via simple bilateral contractual relations, thus providing an example of the view that financial stability is a "public good," likely to be under-provided in equilibrium. We also show that the presence of this externality may manifest itself in equilibrium financial networks that (relative to the socially efficient level) are either too sparsely or too densely connected. The financial network externality and the consequent inefficiency also imply that there is room for welfare-improving government interventions. We leave the analysis of the optimal policy in this class of models to future work.

Finally, we establish a second form of phase transition when long-term returns are partially pledgeable. In particular, we show that if the fraction of the returns that is pledgeable is above a certain threshold, there are no network effects: the excess liquidity of the baking system can be efficiently reallocated to distressed banks and all defaults are avoided, regardless of the structure of the financial network. However, if the pledgeable fraction is below this threshold, network effects cannot be avoided, and the local structure of the financial network plays a defining role in the extent of financial contagion.

There are several areas of future research highlighted by our paper. First, a more complete analysis of network formation in the presence of counterparty risks is an obvious area of investigation. Second, it is important to understand whether and how other forms of financial network interactions (along the lines of those discussed in footnote 7) affect the stability of the system as a whole. Finally and most importantly, a systematic empirical investigation, informed by the theoretical results in this paper, would enable us to measure the key components that play a role in financial stability.

A   Appendix: Proofs

Denote the fraction of obligations of agent $j \neq 0$ to agent i by $q_{ij} = y_{ij}/y_j$. Given that the outside financiers have no debt obligations (i.e., $y_0 = 0$), we define $q_{i0} > 0$ arbitrarily, while assuming that $\sum_{i=1}^{n} q_{i0} = 1$. Let $Q = [q_{ij}] \in \mathbb{R}^{(n+1) \times (n+1)}$ and denote the sum of the payments of bank j to the rest of the banks in the financial network by $x_j = \sum_{i=0}^{n} x_{ij}$. Finally, we define the mapping $\Phi : H \to H$ as

$$\Phi(x) = \left[\min \left\{ Qx + e, y \right\} \right]^+,$$

where $H = \prod_{i=0}^{n} [0, y_i]$. We have the following simple lemma.

Lemma A.1. *For any $x, \hat{x} \in H$,*

$$\|\Phi(x) - \Phi(\hat{x})\|_1 \leq \|Q(x - \hat{x})\|_1. \tag{2}$$

Furthermore, the above inequality is tight only if for all i, either $(Qx+e)_i, (Q\hat{x}+e)_i \in [0, y_i]$ or $(Qx)_i = (Q\hat{x})_i$.

Proof. For any pair of real numbers ξ and $\hat{\xi}$, it is straightforward to verify $|\xi^+ - \hat{\xi}^+| \leq |\xi - \hat{\xi}|$, where the inequality is tight only if $\xi = \hat{\xi}$ or $\xi, \hat{\xi} \geq 0$. Therefore,

$$\left| \Phi_i(x) - \Phi_i(\hat{x}) \right| \leq \left| \min \left\{ (Qx)_i + e_i, y_i \right\} - \min \left\{ (Q\hat{x})_i + e_i, y_i \right\} \right|. \tag{3}$$

On the other hand, it is again easy to verify $|\min\{\xi, \zeta\} - \min\{\hat{\xi}, \zeta\}| \leq |\xi - \hat{\xi}|$ for all real numbers $\xi, \hat{\xi}$ and ζ, where the inequality is tight only if $\xi = \hat{\xi}$ or $\xi, \hat{\xi} \leq \zeta$. As a result,

$$\left| \min \left\{ (Qx)_i + e_i, y_i \right\} - \min \left\{ (Q\hat{x})_i + e_i, y_i \right\} \right| \leq \left| (Qx)_i - (Q\hat{x})_i \right|, \tag{4}$$

for all i. Combining (3) and (4) establishes (2).

To prove the second statement, note that (2) is tight only if inequalities (3) and (4) are tight for all i. Inequality (4) is tight only if either $(Qx)_i = (Q\hat{x})_i$ or $(Qx)_i + e_i, (Q\hat{x})_i + e_i \leq y_i$. On the other hand, inequality (3) is tight only if $(Qx)_i = (Q\hat{x})_i$ or $(Qx)_i + e_i, (Q\hat{x})_i + e_i \geq 0$. Combining these completes the proof. ∎

Proof of Proposition 1

Existence: The mapping $\Phi : H \to H$ continuously maps a convex and compact subset of the Euclidean space to itself. Therefore, by the Brouwer fixed point theorem, there exists $x^* \in H$ such that $\Phi(x^*) = x^*$. This immediately implies that the collection of pairwise interbank payment $\{x_{ij}^*\}$ defined as $x_{ij}^* = q_{ij} x_j^*$ simultaneously solves the collection of equations (1) for all i and j, and thus, corresponds to a payment equilibrium of the financial network.[28]

---

[28] Conversely, any payment equilibrium of the financial network corresponds to a fixed point of Φ.

Generic Uniqueness: Without loss of generality, assume that the financial network is connected.[29] By Lemma A.1, Φ is a non-expansive operator with respect to the ℓ_1-norm. In particular, for any $x, \hat{x} \in H$,

$$\|\Phi(x) - \Phi(\hat{x})\|_1 \quad \leq \quad \|Q(x - \hat{x})\|_1 \tag{5}$$
$$\leq \quad \|Q\|_1 \cdot \|x - \hat{x}\|_1 \tag{6}$$
$$= \quad \|x - \hat{x}\|_1,$$

where the last equality is due to the fact that the column sums of Q are equal to one. Therefore, Φ has two distinct fixed points x and $\hat{x}$ only if all the inequalities above are tight. By Lemma A.1, inequality (5) is tight only if for any given bank i, either $(Qx)_i + e_i, (Q\hat{x})_i + e_i \in [0, y_i]$ or $(Qx)_i = (Q\hat{x})_i$. Denoting the set of banks that satisfy the former by B, we have

$$e_i + \sum_{j=0}^{n} q_{ij} x_j \quad = \quad x_i$$
$$e_i + \sum_{j=0}^{n} q_{ij} \hat{x}_j \quad = \quad \hat{x}_i,$$

for all $i \in B$ with the convention that $e_0 = 0$. Thus,

$$Qx = \begin{bmatrix} x_B - \hat{x}_B \\ 0 \end{bmatrix}.$$

which leads to $\|Qx\|_1 = \|x_B - \hat{x}_B\|_1$. Furthermore, (6) holds as an equality only if $x_i = \hat{x}_i$ for all $i \notin B$. Hence,

$$Q_{BB}(x_B - \hat{x}_B) = x_B - \hat{x}_B, \tag{7}$$

where Q_{BB} is the submatrix of Q corresponding to the banks in B. However, the fact that the financial network is connected (alongside the assumption that $q_{i0} > 0$ for all i) implies that Q is an irreducible matrix and therefore, (7) cannot hold unless B^c is empty.[30] As a consequence,

$$\sum_{i=0}^{n} x_i = \sum_{i=1}^{n} e_i + \sum_{i=0}^{n} \sum_{j=0}^{n} q_{ij} x_j$$

which implies $\sum_{i=1}^{n} e_i = 0$, an equality that holds only for a non-generic set of parameters $z_1, \ldots, z_n$. Thus, the fixed points of Φ, and hence the payment equilibria of the model are generically unique. ∎

Proof of Lemma 1

Denote the set of banks that default on their senior debt by F, the set of banks that default but can meet their senior obligations by D, and the set banks that do not default by $S = (D \cup F)^c$. For any

---

[29]If the financial network is not connected, the proof can be applied to each connected component separately.

[30]An $n \times n$ matrix Q is said to be *reducible*, if for some permutation matrix P, the matrix $P'QP$ is block upper triangular. If a square matrix is not reducible, it is said to be *irreducible*. If Q is a non-negative irreducible matrix with unit column sums, then all eigenvalues of any square submatrix of Q, say $\tilde{Q}$, lie within the unit circle, which implies that $\tilde{Q}x = x$ has no solution. For more on this, see e.g., Berman and Plemmons (1979).

bank $i \in F$, we have

$$\pi_i + T_i = z_i + \sum_{j \neq i} x_{ij},$$

whereas for $i \in D$,

$$\pi_i + T_i = v.$$

On the other hand, for any bank $i \in S$ which does not default,

$$\pi_i + T_i = A + z_i - y + \sum_{j \neq i} x_{ij}.$$

Finally, note that in any regular financial network, the outside financiers do not lend to the banks and instead, invest their capital in a linear technology with return r, implying that $\pi_0 = 0$. Therefore, the social surplus is given by

$$
\begin{aligned}
u &= (A - y)|S| + v|D| + \sum_{i \notin D} z_i + \sum_{i \notin D} \sum_{j \neq i} x_{ij} \\
&= (A - y)|S| + \sum_{i=1}^{n} z_i + \sum_{i \in S} \sum_{j \neq i} y_{ij}
\end{aligned}
$$

where the second equality is due to the fact that for $i \in D$ we have, $\sum_{j \neq i}(x_{ji} - x_{ij}) = z_i - v$. Hence, $u = na - m\epsilon + A(n - \sharp\text{defaults})$. ∎

Two Auxiliary Lemmas

Lemma A.2. *The number of bank defaults satisfies*

$$m \leq \sharp(\text{defaults}) < \frac{m\epsilon}{a - v}$$

where m is the number of realizations of negative shocks in the network.

Proof. Given that the total interbank liabilities of each bank is equal to its total interbank claims and that $v > a - \epsilon$, any bank that is hit with a negative shock defaults. Hence, the lower bound is trivial. To obtain the upper bound, note that for any bank i that defaults but can meet its senior obligations in full, we have

$$z_i + \sum_{j \neq i} x_{ij} = v + \sum_{j \neq i} x_{ji}.$$

Denoting the set of such banks by D and summing over all $i \in D$ imply

$$\sum_{i \in D} z_i + \sum_{i \in D} \sum_{j \neq i} x_{ij} = v|D| + \sum_{i \in D} \sum_{j \neq i} x_{ji}. \tag{8}$$

On the other hand, for any bank i that defaults on its senior obligations (if such a bank exists), we have

$$\sum_{j \neq i} x_{ij} + z_i < v.$$

28

Summing over the set of all such banks, F, implies

$$\sum_{i \in F} \sum_{j \neq i} x_{ij} + \sum_{i \in F} z_i < v|F|. \tag{9}$$

Adding (8) and (9) leads to

$$m\epsilon - (a - v)\sharp(\text{defaults}) \geq \sum_{j \notin S} \sum_{i \in S} (y_{ij} - x_{ij}),$$

where S is the set of banks that do not default. By definition, the right-hand side of the above equality is strictly positive, proving that the number of defaults is strictly smaller than $m\epsilon/(a - v)$. ∎

Lemma A.3. *If $\epsilon < \epsilon_m^*$, then at least one bank does not default, where $\epsilon_m^* = n(a - v)/m$. On the other hand, if $\epsilon > \epsilon_m^*$, then at least one bank defaults on its senior obligations.*

Proof. Suppose that $\epsilon < \epsilon_m^*$ and that all banks default. Therefore,

$$z_i + \sum_{j \neq i} x_{ij} \leq v + \sum_{j \neq i} x_{ji},$$

for all banks i. Summing over i implies

$$na - m\epsilon \leq nv,$$

which contradicts the assumption that $\epsilon < \epsilon_m^*$.

To prove the second statement, suppose that $\epsilon > \epsilon_m^*$ and that no bank defaults on its senior obligations. Thus,

$$z_i + \sum_{j \neq i} x_{ij} \geq v + \sum_{j \neq i} x_{ji},$$

for all banks i. Summing over i implies $na - m\epsilon \geq nv$, leading to a contradiction. ∎

Proof of Proposition 2

Proof of part (a). Without loss of generality assume that bank 1 is hit with the negative shock. By Lemma A.3, bank n does not default as it is the bank furthest away from the distressed bank. Moreover, as long as $y > y^* = (n-1)(a-v)$, bank 1 — and consequently all banks — can meet their senior obligations in full. This is due to the fact that $y + a - \epsilon > v$. Given that the set of banks that default form a connected chain, say of length τ, the repayment of the last bank in default to its sole creditor satisfies

$$x_{\tau+1,\tau} = y + \tau(a - v) - \epsilon.$$

On the other hand, given that bank $\tau + 1$ does not default, we have

$$a + x_{\tau+1,\tau} > y + v.$$

As a result, $\tau \geq \epsilon/(a-v) - 1$, implying that the number of defaults reaches the upper bound established in Lemma A.2. Hence, the ring network is the least stable and least resilient financial network.

Proof of part (b). By Lemma A.3, as long as $\epsilon < \epsilon^*$, there exists at least one bank that does not default. Given the full symmetry in the complete network, the $n-1$ banks that are not hit with the negative shock can thus meet their obligations in full. Hence, the complete financial network is the most stable and most resilient regular financial network.

Proof of part (c). Consider the financial network constructed as the γ-convex combination of the ring and the complete financial networks. Without loss of generality, we assume that bank 1 is hit with the negative shock. Define γ_d to be the value at which banks 1 through $d-1$ default while bank d is at the verge of default. At this value of γ, we have

$$x_1 = \left(\frac{1-\gamma_d}{n-1}\right)[(x_1 + \cdots + x_d) - x_1 + (n-d)y] + \gamma_d y + (a - v - \epsilon) \tag{10}$$

$$x_i = \left(\frac{1-\gamma_d}{n-1}\right)[(x_1 + \cdots + x_d) - x_i + (n-d)y] + \gamma_d x_{i-1} + (a - v) \tag{11}$$

for $2 \leq i \leq d$. Hence,

$$\Delta_2 = x_2 - x_1 = \gamma_d \Delta_1 - \left(\frac{1-\gamma_d}{n-1}\right)\Delta_2 + \epsilon$$

$$\Delta_i = x_i - x_{i-1} = \gamma_d \Delta_{i-1} - \left(\frac{1-\gamma_d}{n-1}\right)\Delta_i,$$

where $\Delta_1 = x_1 - y$. Thus, for all $2 \leq i \leq d$,

$$\Delta_i = \beta^{i-1}(\Delta_1 + \epsilon/\gamma_d),$$

where for notational simplicity we have defined $\beta = (1 + \frac{1-\gamma_d}{n-1})^{-1}\gamma_d$. Given that $x_d = y$, the terms Δ_i must add up to zero, that is,

$$\sum_{i=1}^{d} \Delta_i = \left(\frac{1-\beta^d}{1-\beta}\right)\Delta_1 + \left(\frac{1-\beta^{d-1}}{1-\beta}\right)\left(\frac{\beta\epsilon}{\gamma_d}\right) = 0,$$

which immediately implies

$$\Delta_1 = -\left(\frac{1-\beta^{d-1}}{1-\beta^d}\right)\left(\frac{\beta\epsilon}{\gamma_d}\right).$$

Therefore,

$$x_i = y + \sum_{s=1}^{i} \Delta_s = y + \left(\frac{\beta^{d-1} - \beta^{i-1}}{1-\beta^d}\right)\left(\frac{\beta\epsilon}{\gamma_d}\right),$$

and as a result,

$$\sum_{i=1}^{d-1} x_i = (d-1)y + \left[\frac{d\beta^{d-1}(1-\beta) - (1-\beta^d)}{\gamma_d(1-\beta)(1-\beta^d)}\right]\beta\epsilon.$$

30

On the other hand, from (10) and (11), we have

$$\sum_{i=1}^{d-1} x_i = (d-1)y + \frac{d(a-v) - \epsilon}{(n-d)(1-\gamma_d)/(n-1)}.$$

Equating the above two equalities thus leads to

$$n(a-v)/\epsilon - 1 = \left[\frac{\beta^{d-1}(1-\beta)}{1-\beta^d}\right](n-d).$$

Therefore, the value γ_d at which $d-1$ banks default and bank d is at the verge of default must satisfy the above equality. For a fixed value of d, the right-hand side is increasing in β (and hence in γ_d).[31] As a consequence, in order for the right-hand side to remain equal to the constant on the left-hand side, d has to decrease with γ. Thus, there will be weakly more defaults as γ increases. ∎

Proof of Proposition 3

Proof of part (a). First consider the complete financial network. By Lemma A.3, the distressed bank defaults on its senior obligations. Suppose that the remaining $n-1$ banks do not default and that they can all meet their obligations in full.[32] This would be the case only if

$$(n-2)\frac{y}{n-1} + (a-v) \geq y,$$

where we are using the fact that the distressed bank does not pay anything to its junior creditors. The above inequality, however, contradicts the assumption that $y > y^*$. Hence, all banks default, implying that the complete network is the least resilient and the least stable financial network.

Now consider the ring financial network and assume without loss of generality that bank 1 is hit with the negative shock. Once again by Lemma A.3, bank 1 defaults on its senior obligations and hence, does not pay anything to its more junior creditor, bank 2. Thus, if some bank does not default, the length of the default cascade, τ, must satisfy

$$\tau(a-v) > y,$$

which in light of the assumption that $y > y^*$ guarantees that $\tau > n-1$, leading to a contradiction. Hence, all banks default, implying that the ring financial network is the least stable and resilient financial network.

Proof of part (b). By definition, any regular δ-connected financial network contains at least one δ-component M. Given that the total obligations of banks in M^c to any bank $i \in M$ is at most δ, the total obligations of banks in M to i is at least $y - \delta$; that is, $\sum_{j \in M} y_{ij} > y - \delta$.

---

[31]This can be easily verified by noticing that $\beta^{d-1}(1-\beta) = (1-\beta^d)\left(1 + \beta^{-1} + \cdots + \beta^{-(d-1)}\right)^{-1}$.

[32]Note that given the symmetric structure of the complete network and the uniqueness of payment equilibrium, either all other banks default together or they all meet their obligations fully at the same time.

If the negative shock hits a bank outside of M, all banks in M can meet their obligations in full as long as

$$y < a - v + \min_{i \in M}\left\{\sum_{j \in M} y_{ij}\right\},$$

which holds whenever $\delta < a - v$. A similar argument shows that when the shock hits a bank in M, all banks in M^c can meet their obligations in full. Thus, the financial network is strictly more stable and resilient than both the complete and the ring networks. ∎

Proof of Proposition 4

Proof of part (a). First consider the complete financial network. If $\epsilon < \epsilon_m^*$, at least one bank does not default. Given the symmetry, all $n - m$ banks that are not hit with a negative shock do not default either, implying that the complete network is the most stable and resilient financial network in the face of small shocks.

Now consider the ring financial network and assume that m consecutive banks, labeled $i + 1$ through $j = i + m$, are hit with negative shocks. An immediate observation is that all banks in default also form a connected chain, say of length $\tau \geq m$, the last of which is labeled $s = i + \tau$. In view of Lemma A.3, bank i does not default, as it is the bank furthest away from the realized shocks. As a result, as long as $y > y_m^* = (n - m)(a - v)$, in the unique payment equilibrium of the financial network, all banks can meet their senior obligations v in full. This can be established by verifying that bank j — which is the bank facing the most amount of potential distress — can pay its senior debts. In particular,

$$x_{j,j-1} = y + (m - 1)(a - \epsilon - v),$$

guaranteeing that $x_{j,j-1} + a - \epsilon > v$. Given that all banks can meet their senior obligations, we have

$$x_{s+1,s} = y + \tau(a - v) - m\epsilon$$

where $s = i + \tau$ is the index of the last bank on the chain that defaults. On the other hand, given that $s + 1$ does not default, we have $y \leq a - v + x_{s+1,s}$. As a result,

$$\tau = \left\lceil \frac{m\epsilon}{a - v} \right\rceil - 1 \geq \frac{m\epsilon}{a - v} - 1. \tag{12}$$

Hence, when shocks hit m consecutive banks on the credit chain, the number of bank failures reaches the upper bound established by Lemma A.2, implying that the ring network is the least resilient financial network.

Proof of part (b). The proof follows a logic similar to that of Proposition 3. We first prove that if $\epsilon > \epsilon_m^*$, then the complete network is the least stable and resilient financial network. In particular, we

show that all banks default. By Lemma A.3, the m distressed banks default on their senior obligations. The remaining $n - m$ banks do not default only if

$$(n - m - 1)\frac{y}{n - 1} + (a - v) \geq y.$$

The above inequality, however, can hold only if $y < \hat{y}_m = (n-1)(a-v)/m \leq y_m^*$. Hence, the complete network is the least resilient and the least stable financial network as all n banks default.

We next show that if $\epsilon > \epsilon_m^*$, then all n banks in the ring network fail as well. Suppose not, and that there exists a bank that can pay all its creditors in full. On the other hand, by Lemma A.3, there is also a bank that defaults on its senior obligations v. Consider the path on the ring network connecting bank j to bank l, such that (i) j defaults on its senior debt; (ii) l pays all its creditors in full; and (iii) all banks on the path default but can pay back their senior debt. Denote the length of the path connecting j to l by τ (see Figure 6), and suppose that there are h negative shocks realized on this path.
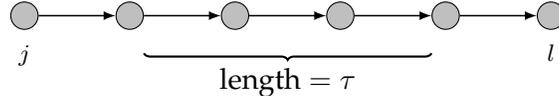

Figure 6: There are τ banks connecting j to l, all of which default, but can meet their senior obligations.

Given that j does not pay anything to its junior creditor (which is the first bank on the path connecting it to l) and that l does not default, we have $(\tau + 1)(a - v) - h\epsilon \geq y$, implying that

$$\begin{aligned} \tau &> \frac{y_m^* + h\epsilon_m^*}{a - v} - 1 \\ &= n - m - 1 + \frac{hn}{m}. \end{aligned}$$

On the other hand, the remaining $m - h$ shocks hit banks that are not on the path connecting j to l. Thus, the total number of defaults is at least $\tau + m - h$, implying

$$\sharp\text{defaults} > n - 1 + h\left(\frac{n}{m} - 1\right) \geq n - 1.$$

This, however, contradicts the assumption that at least one bank does not default.

Finally, consider a δ-connected financial network, containing a δ-component M. Given that the total obligations of banks in M^c to any bank $i \in M$ is at most δ, the total obligations of banks in M to i is at least $y - \delta$; that is, $\sum_{j \in M} y_{ij} > y - \delta$ for all i. If the negative shock hits a bank outside of M and as long as $y < a - v + \min_{i \in M}\left\{\sum_{j \in M} y_{ij}\right\}$, all banks in M can meet their obligations in full. Thus, as long as $\delta < a - v$, the financial network is strictly more stable than both the complete and the ring networks.

Proof of part (c). An argument similar to the one invoked in the proof of part (b) shows that the when $\epsilon > \epsilon_m^*$ and $y > \hat{y}_m$, all banks in the complete network default. Therefore, the complete network is the least stable and resilient financial network.

To prove that the ring financial network is more stable than the complete network, we show that there exists a realization of the shocks for which at least one bank in the ring network does not default. In particular, consider the situation in which m consecutive banks, labeled 1 through m, are hit with negative shocks. By Lemma A.3, in the unique payment equilibrium, bank m defaults on its senior debt. Therefore, the length of the cascade of defaults following bank m, denoted by τ, satisfies

$$\tau(a - v) < y \leq (\tau + 1)(a - v).$$

Thus, the number of defaults in the whole network is

$$\begin{aligned}
\sharp\text{defaults} &= m + \tau \\
&< m + \frac{y_m^*}{a - v} \\
&= n,
\end{aligned}$$

implying that at least one bank does not default. Hence, the ring network is strictly more stable than the complete network. ∎

Proof of Proposition 5

We first prove that the 3-chain financial network is efficient if and only if

$$(r - 1)k \geq pA. \tag{13}$$

Given that there are only two potential interbank bilateral contracts, there are only four possible financial networks that can arise. It is easy to show that if banks 2 and 3 lend to banks 1 and 2, respectively, the social surplus is equal to

$$u_{3\text{-chain}} = 2a - p\epsilon + k + 2(1 - p)A,$$

where $2(1 - p)A$ captures the fact that a negative shock to bank 1 would lead to the default of (and hence, costly liquidation by) banks 1 and 2. Furthermore,

$$\begin{aligned}
u_{23} &= 2a - p\epsilon + 2k - rk + (2 - p)A \\
u_{12} &= 2a - p\epsilon + 2k - rk + 2(1 - p)A \\
u_{\text{empty}} &= 2a - p\epsilon + 3k - 2rk + (2 - p)A,
\end{aligned}$$

where u_{ij} refers to the social surplus in the financial network in which only bank j lends to bank i, and u_{empty} is the social surplus in the financial network with no interbank borrowings. The above equalities immediately imply that the 3-chain financial network is efficient if and only if (13) holds.

We next show that (13) is a necessary condition for the 3-chain to be an equilibrium financial network. Suppose that the 3-chain financial network along with the collection of posted contracts

$(\mathbf{R}_0, \mathbf{R}_2, \mathbf{R}_3)$ correspond to an equilibrium.[33] Given that the outside financiers are competitive and have to break even, the contract posted by the representative outside financier, $\mathbf{R}_0 = (R_{01}, (R_{02}, R'_{02}))$, satisfies

$$
\begin{aligned}
rk &= (1-p)R_{01}k + p(a - v - \epsilon + k) & (14) \\
rk &= (1-p)R_{02}k + p(2(a - v) - \epsilon + k) & (15) \\
rk &= R'_{02}k,
\end{aligned}
$$

where R_{02} and R'_{02} are interest rates faced by bank 2 contingent on its decision whether to lend to bank 1 or not, respectively. It is also immediate to verify that the equilibrium interest rate offered by bank 2 to bank 1 has to be equal to the one offered by the outside financiers; that is, $\mathbf{R}_2 = R_{01}$. Else, either bank 1 would have borrowed from the outside financiers, or bank 2 could have increased its profits by posting a higher interest rate. A similar argument establishes that, as long as bank 2 lends to bank 1, bank 3 would not charge an interest rate different than what is offered by the outside financiers. In other words, $\mathbf{R}_3 = (R_{02}, R'_{32})$, where R'_{32} is the interest rate at which bank 2 can borrow from bank 3 if it decides not to lend to bank 1. Thus, given the equilibrium interest rates, the profits of the banks are equal to[34]

$$
\begin{aligned}
\pi_1 &= a - v - (r-1)k - p\epsilon + (1-p)A \\
\pi_2 &= a - v + (1-p)A \\
\pi_3 &= rk.
\end{aligned}
$$

Finally, for the 3-chain financial network to be part of an equilibrium, bank 3 should not be able to make a strictly higher profit by deviating and posting a new contract $\tilde{\mathbf{R}}_3 = (\tilde{R}_{32}, \tilde{R}'_{32})$. In particular, there should not exist an interest rate $\tilde{R}'_{32}$ for which both banks 2 and 3 would make strictly higher profits if bank 2 hoards cash. Given that the profits of the banks when bank 2 hoards cash are

$$
\begin{aligned}
\tilde{\pi}_2 &= a - v + k - \tilde{R}'_{32}k + A \\
\tilde{\pi}_3 &= \tilde{R}'_{32}k,
\end{aligned}
$$

such a deviation would be profitable if there exists $\tilde{R}'_{32}$ such that

$$
k + pA > \tilde{R}'_{32}k > rk
$$

as it would guarantee that $\tilde{\pi}_2 > \pi_2$ and $\tilde{\pi}_3 > \pi_3$. Thus, if (13) does not hold, the 3-chain financial network cannot be part of an equilibrium.

The proof is complete once we show that (13) is also a sufficient condition for the 3-chain financial network to be part of an equilibrium. In particular, we show that if (13) holds, then the collection

---

[33]Note that since bank 1 has no lending opportunity, the contract $\mathbf{R}_1$ is not part of the equilibrium definition.

[34]Note that given that hoarding cash by bank 2 is off the equilibrium path, the equilibrium profits of the banks are independent of the value R'_{32}.

of contracts $\mathbf{R}_0 = (R_{01}, (R_{02}, r))$, $\mathbf{R}_2 = R_{01}$, and $\mathbf{R}_3 = (R_{02}, r)$ and the 3-chain financial network constitute an equilibrium, where R_{01} and R_{02} solve equations (14) and (15), respectively. To this end, first take the posted contracts as given and consider the subgame that follows. Banks 1 and 2 have no unilateral incentives to borrow from the outside financiers, as they would be facing the same interest rates offered by banks 2 and 3, respectively. Furthermore, bank 2 does not have an incentive to deny lending to bank 1. In particular, if bank 2 hoards cash, it would make a profit of $\hat{\pi}_2 = a - v + k + A - rk$ which is strictly smaller than $\pi_2 = a - v + (1 - p)A$. Finally, note that bank 3 does not have an incentive to hoard either. Because, if bank 3 hoards cash, its profit would be equal to $\hat{\pi}_3 = k$, which is strictly smaller than its profits in the 3-chain financial network. Thus, given $(\mathbf{R}_0, \mathbf{R}_2, \mathbf{R}_3)$, the 3-chain financial network is a Nash equilibrium of the subgame that follows the posting of the contracts.

Finally, we show that no bank has an incentive to unilaterally deviate by posting a different contract. Given that $\mathbf{R}_0 = (R_{01}, (R_{02}, r))$, bank 2 has no incentive to post anything other than $\mathbf{R}_2 = R_{01}$. Similarly, bank 3 cannot make strictly positive profits by posting a contract that offers bank 2 an interest rate other than R_{02} if bank 2 does not hoard cash. Thus, the only possible profitable unilateral deviation would be for bank 3 to post a contract $\tilde{\mathbf{R}}_3 = (R_{02}, \tilde{R}'_{32})$ in which $\tilde{R}'_{32} < r$. However, as we showed earlier, as long as (13) holds, such a deviation either (i) is not profitable to bank 3, or (ii) would not induce bank 2 to hoard cash. Thus, the specified collection of contract and the 3-chain financial network constitute a subgame perfect Nash equilibrium. ∎

Proof of Proposition 6

Proof of part (a). Define $\tau = \lceil \epsilon/(a - v) \rceil - 1$ as the extent of financial contagion in the ring financial network and let $\underline{\alpha} = 2(\tau - 1)/(n - 1)$. We show that if $\underline{\alpha}A < (r - 1)k$, a collection of contracts of the form $\mathbf{R}_i = (R, R')$ for all $i \in \{0, 1, \ldots, n\}$ along with the ring financial network constitute a symmetric equilibrium, where R and R' are the interest rates at which the lender is willing to lend to the borrower contingent on whether the latter lends or hoards cash, respectively. In particular, we choose the pair (R, R') to satisfy the indifference equations of the outside financiers:

$$rk = \left(\frac{n - \tau}{n}\right) Rk + \frac{1}{n} \sum_{s=1}^{\tau} \left[s(a - v) - \epsilon + Rk \right] \tag{16}$$

$$rk = \left(\frac{n - 1}{n}\right) R'k + \frac{1}{n}(a - v - \epsilon + k). \tag{17}$$

We first take the pair of interest rates (R, R') as given, and consider the subsequence subgame. Note that bank i has an incentive to deny lending to bank $i - 1$ and hoard cash instead. In particular, if bank i hoards cash, it would face the interest rate R' and as a result, its profits would be equal to

$$\tilde{\pi} = a - v + \left(\frac{n - 1}{n}\right) A - \frac{\epsilon}{n} - (r - 1)k.$$

On the other hand, if bank i does not hoard cash, it would make a profit equal to

$$\pi_{\text{ring}} = \left(\frac{n - \tau}{n}\right) A + a - v - \frac{\epsilon}{n}, \tag{18}$$

which is strictly larger that $\tilde{\pi}$. Furthermore, bank i has no incentive to deviate in its borrowing behavior either, as it would face interest rate R no matter if it borrows from bank $i + 1$ or the outside financiers. Therefore, given the collection of contracts $\mathbf{R}_i = (R, R')$, the ring financial network is a Nash equilibrium of the corresponding subgame.

It is thus sufficient to show that no bank has an incentive to deviate by posting a different contract. Clearly, no bank can make a strictly higher profit by offering an interest rate other than R for the case that its designated borrower does not hoard cash. Therefore, the only possible profitable deviation is for a bank i to post $\tilde{\mathbf{R}}_i = (R, \tilde{R}')$ where $\tilde{R}' < R'$. For such a deviation to be profitable, however, there should exist an interest rate $\tilde{R}'$ such that $\tilde{\pi}_i > \pi_{\text{ring}}$ and $\tilde{\pi}_{i+1} > \pi_{\text{ring}}$ where,

$$\tilde{\pi}_i = \left(\frac{n-1}{n}\right)(a - v + k - \tilde{R}'k + A)$$

$$\tilde{\pi}_{i+1} = \left(\frac{n-2}{n}\right)(a - v + \tilde{R}'k - Rk + A).$$

Thus, such an interest rate exists only if

$$k - \frac{1}{n-1}(a - v - \epsilon) + \left(\frac{\tau - 1}{n - 1}\right)A > Rk + \frac{1}{n-2}(2(a - v) - \epsilon) - \left(\frac{\tau - 2}{n - 2}\right)A,$$

which implies

$$\left(\frac{\tau - 1}{n - 1} + \frac{\tau - 2}{n - 2}\right)A + \left(\frac{1}{n-1} + \frac{1}{n-2}\right)\epsilon > (r - 1)k + \left(\frac{\tau - 1}{2n} + \frac{1}{\tau(n-1)} + \frac{2}{\tau(n-2)}\right)\epsilon,$$

where we are using the fact that $\tau = \lceil \epsilon/(a - v) \rceil - 1$. The above inequality, however, contradicts the assumption that $\underline{\alpha}A < (r - 1)k$. Therefore, such a deviation would not be profitable for by i, completing the proof.

Proof of part (b). Given that the ring financial network is symmetric, there are exactly τ defaults in the presence of a single negative shock. Therefore, the social surplus is equal to

$$u_{\text{ring}} = na - \epsilon + (n - \tau)A.$$

If, on the other hand, the social planner forces a single bank i to hoard cash instead of lending to bank $i - 1$, (that is, by setting $\ell_{i,i-1} = 0$ and $\ell_{j,j-1} = k$ for $j \neq i$) the social surplus would be equal to

$$u_{\text{hoard}} = na - \epsilon + (n - \mathbb{E}[\tau_{\text{hoard}}])A - (r - 1)k,$$

where $\mathbb{E}[\tau_{\text{hoard}}]$ is the expected number of defaults. Given that bank i is no longer at the risk of default due to contagion, the extent of contagion would be strictly smaller than τ if any of the banks indexed $i - \tau + 1$ through $i - 1$ are hit by the negative shock, whereas there would be a longer cascade of defaults, say of length $\tau' \geq \tau$, if bank i itself is distressed. This is due to the fact that hoarding cash

implies a smaller return on i's capital, and hence, a smaller cushion to absorb the shock at $t = 1$. Hence,

$$\mathbb{E}[\tau_{\text{hoard}}] = \frac{1}{n}\left[(n - \tau - 1)\tau + \sum_{s=1}^{\tau} s + \tau'\right]$$

$$\leq \tau + 1 - \frac{\tau(\tau + 1)}{2n},$$

where the inequality is due to the fact that $\tau' \leq n$. Therefore, as long as

$$(r - 1)k < \left(\frac{\tau(\tau + 1)}{2n} - 1\right)A,$$

the social planner can increase the social surplus by forcing bank i not to lend to bank $i - 1$. ∎

Proof of Proposition 7

Proof of part (b). We establish the inefficiency of the double-ring financial network by showing that the social surplus is higher in the interlinked rings financial network. Given that there are τ defaults in the double-ring financial network, the social surplus is equal to

$$u_{\text{2-ring}} = na - \epsilon + (n - \tau)A.$$

On the other hand, the social surplus in the interlinked rings network is

$$u_{\text{linked}} = na - \epsilon + (n - \hat{\tau})A,$$

where $\hat{\tau}$ is the number of defaults. Thus, it is sufficient to show that $\hat{\tau} < \tau$. Suppose without loss of generality that bank n in the interlinked financial network is hit with the negative shock. Hence, if banks $2q - 1$ and $2q$ default, their total repayments on their obligations are equal to

$$x_{2q-1} = x_{2q} = Rk + (q + 1/2)(a - v) - \epsilon/2.$$

As a result, the total number of defaults in the interlinked rings financial network would be equal to $2s + 1$, where s is the largest integer for which $x_{2s} < Rk$. Hence,

$$\hat{\tau} = 2\left\lceil \frac{\epsilon}{2(a - v)} - 3/2 \right\rceil + 1.$$

Given that $\epsilon/(a - v) \leq \tau + 1$, we have

$$\hat{\tau} \leq 2\lceil \tau/2 - 1 \rceil + 1 = 2\lceil \tau/2 \rceil - 1,$$

which implies that if τ is even, then $\hat{\tau} < \tau$. Hence, the double-ring financial network is socially inefficient.

Proof of part (a). We show that a collection of contracts of the form $\mathbf{R} = (R, R')$ for all $i \in \{0, 1, \dots, n\}$ along with the double-ring financial network constitute a symmetric equilibrium, where R and R' are the interest rates at which bank i is willing to lend to a borrower contingent on whether the latter lends to others (regardless of whether it lends to one or two other banks) or hoards cash, respectively. We choose the pair (R, R') to satisfy the indifference equations of the outside financiers, determined by (16) and (17).

Given that in the double-ring financial network each bank effectively belongs to a single credit chain, the results established in Proposition 6 are applicable. In particular, the double-ring financial network is a Nash equilibrium of the subgame that follows the posting of contract $\mathbf{R}$ by all agents. Furthermore, no bank has an incentive to deviate and charge an interest rate other than R or R' when its borrower lends fully or hoards cash, respectively. Therefore, to verify that $\mathbf{R}$ and the double-ring financial network constitute an equilibrium, it is sufficient to show that no lender bank $i + 2$ has an incentive to charge an interest rate that would lead the borrower bank i to split its lending between two counterparties. We establish this statement by showing that if the interest rate is kept at R, both banks i and $i + 2$ would make strictly smaller profits when bank i splits its lending. Hence, there is no interest rate at which (i) bank $i + 2$ can induce bank i to split its lending; and (ii) bank $i + 2$ is strictly better off than in the double-ring financial network.

Recall from (18) that the profit of a bank in a credit chain is equal to

$$\pi_{\text{ring}} = \left(\frac{n - \tau}{n}\right) A + a - v - \frac{\epsilon}{n}.$$

On the other hand, if bank i deviates and splits its lending between the two borrower banks (i.e., banks $i - 2$ and $i - 1$ assuming that i is odd), then it would default on its obligations to bank $i + 2$ if any bank j satisfying $i - 2\tau + 4 \leq j \leq i$ is hit with the negative shock. Therefore, if the interest rates are kept constant at R,

$$\hat{\pi}_i = \left(\frac{n - 2\tau + 3}{n}\right) A + \left(\frac{n - \tau + 2}{n}\right)(a - v) - \frac{\epsilon}{n}.$$

A similar argument implies that bank $i + 2$ defaults with probability $2\tau - 6$ and hence,

$$\hat{\pi}_{i+2} = \left(\frac{n - 2\tau + 6}{n}\right) A + \left(\frac{n - \tau + 5}{n}\right)(a - v) - \frac{\epsilon}{n}.$$

The above equations immediately imply that if $\tau > 3$, then $\hat{\pi}_i, \hat{\pi}_{i+2} < \pi_{\text{ring}}$, completing the proof. ∎

Proof of Proposition 8

We first state and prove a simple lemma. Suppose that $p = 0$ and define R_c as the interest rate at which the outside financiers break even in a complete financial network when all banks are also lending at rate R_c. In other words, R_c solves

$$rk = \left(\frac{n - 1}{n}\right) R_c k + \frac{1}{n}(a - v - \epsilon_\ell + R_c k).$$

Furthermore, for $0 \leq s \leq n-1$, define R_s as the interest rate at which the outside financiers are willing to lend to a bank i who deviates from the complete network by only lending to s other banks.[35]

Lemma A.4. *Suppose that $p = 0$. Then, $R_s > R_c$ for all $s < n - 1$.*

Proof. Suppose that bank i lends to s other banks, where $s < n - 1$. Also suppose that the outside financiers break even by charging bank i an interest rate of $R_s \leq R_c$. It is easy to verify that bank i defaults on its obligations with probability at least $1/n$. In particular, even if $R_s = r$, the bank cannot meet all its obligations, as $a - \epsilon_\ell + R_c k < v + rk$. Furthermore, given that bank i is lending at rate R_c, the total default repayment of i on its debt to the outside financiers is strictly less than $R_c k + a - v - \epsilon_\ell$. Hence, the expected repayment of bank i on its debt to the outside financiers satisfies

$$\mathbb{E}[x_{0i}] < \left(\frac{n-1}{n}\right) R_s k + \frac{1}{n}\left(a - v - \epsilon_\ell + R_c k\right) \leq rk,$$

which is a contradiction. ∎

Proof of part (a). Given that the shock is of size of ϵ_ℓ with probability one, the social surplus in the economy is equal to

$$u = na - \epsilon_\ell + (n - \mathbb{E}[\tau]) A - (c_1 + \cdots + c_n)(r - 1),$$

where c_i is the amount of cash hoarded by bank i and $\mathbb{E}[\tau]$ is the expected number of defaults in the financial network. As the above equality suggests, only the last two terms depend on the structure of the financial network. Given that $r > 1$, the last term achieves its maximum value of zero when no bank hoards cash. Clearly, the complete network satisfies this condition. Furthermore, by Proposition 2, there is exactly one failure in the complete network regardless of the realization of the shock, and hence, $\mathbb{E}[\tau] = 1$. Thus, the proof is complete once we show that the expected number of failures in all symmetric financial networks is at least one. Suppose not. This implies that there exists a realization of the shocks for which no bank fails. By symmetry, however, this means that no bank fails in the financial network, and as a result, all banks can borrow at the risk-free interest rate r. Therefore, the total amount owed to any given bank i is less than or equal to its total debt of value rk. This however, implies that the bank defaults if it is hit with the negative shock, which is a contradiction. Thus, the complete network is the most efficient symmetric financial network.

Proof of part (b). It is sufficient to construct a symmetric financial network for which the social surplus is larger than that of the complete network. First, consider the complete financial network. Conditional on the realization of ϵ_ℓ and ϵ_h, there are 1 and n defaults, respectively. Therefore, the social surplus in the economy is

$$u_{\text{complete}} = na + (1 - p)(n - 1)A - [(1 - p)\epsilon_\ell + p\epsilon_h].$$

---

[35]Note that by definition, $R_{n-1} = R_c$.

Next, consider the empty financial network in which banks do not lend to one another at all and instead, hoard all their cash. In this case, the social surplus in the economy is

$$u_{\text{empty}} = na + (n-1)A - [(1-p)\epsilon_\ell + p\epsilon_h] - nk(r-1),$$

where we are using the fact that a bank defaults only if it is hit with (either a small or a large) shock. Thus, for any given p and n, the complete network is inefficient if

$$p(n-1)A \geq n(r-1)k,$$

which completes the proof.

Proof of part (c). We first show that if $p = 0$, then the complete financial network is part of an equilibrium. In particular, we show that if all banks and the outside financiers post the contract $\mathbf{R} = (R_0, \ldots, R_{n-1})$, then the complete financial network can be supported as part of an equilibrium, where R_s is defined above.

As a first step, we show that given $\mathbf{R}$, the complete network is a Nash equilibrium of the subsequent subgame. Given that all banks and the outside financiers are posting the same contract, no bank has an incentive to deviate and borrow from the outside financiers instead of other banks. Furthermore, no bank has an incentive to deviate and lend less either. By Lemma A.4, such a deviation implies that the bank would face a higher interest rate from its creditors, while getting a smaller rate of return on its own investment. Therefore, given the contracts, no bank has a profitable unilateral deviation.

Next, we show that no bank has an incentive to deviate by posting a contract other than $\mathbf{R}$. Clearly, a given bank cannot make strictly higher profits by charging an interest rate other than R_c for the case that its borrowers are not hoarding cash. In particular, if the bank chargers an interest rate higher than R_c, the borrower banks would switch to the outside financiers. Thus, any potential profitable deviation should involve changing the interest rates that a bank is willing to charge contingent on the borrowers hoarding cash. Note that such deviations would lead to different payoffs only if the banks do end up hoarding cash in the subsequent subgame. Thus, suppose that a bank i deviates by posting a contract $\hat{\mathbf{R}} \neq \mathbf{R}$ and that at least one bank, say bank j, hoards cash (partially or fully) in the subgame that follows. Now it is easy to verify that the corresponding element of the vector of interest rates $\hat{\mathbf{R}}$ cannot be larger than R_c. In particular, if that is the case, then j can simply stop hoarding cash, and hence, face the smaller interest rate of R_c instead. On the other hand, no element of $\hat{\mathbf{R}}$ can be smaller than or equal to R_c either. In particular, if bank i charges an interest rate that is smaller than or equal to R_c, it would be receiving a strictly smaller repayment on its loan to j, regardless of whether j is defaulting or not. Hence, no bank has a profitable unilateral deviation in posting a contract other than $\mathbf{R}$.

To summarize, we have so far shown that when $p = 0$, the complete financial network is part of an equilibrium. More specifically, we showed that any deviation that would lead to a financial network

41

structure other than the complete network in the subsequent subgame, would lead to a strictly lower profit for the deviating bank. Therefore, by continuity, there exists a small enough $\bar{p}$ such that for $p < \bar{p}$, the complete network is part of a subgame perfect Nash equilibrium of the game. ∎

Proof of Proposition 9

If $\epsilon > \epsilon^*$, then there is simply not enough liquidity in the system to forestall all defaults, regardless of the value of κ. Therefore, throughout the proof, we assume that $\epsilon < \epsilon^*$.

Proof of part (a). The argument relies on transforming the pattern of interbank renegotiations into a network flow problem and invoking the maximum-flow minimum-cut theorem of Ford and Fulkerson (1956).[36] In particular, consider the multi-source multi-destination network flow problem defined as follows: (i) each bank i for which $z_i = a$ is a source with value of $a - v$, (ii) each distressed bank is a sink of value $\epsilon + v - a$, and finally, (iii) each vertex in the network has a capacity of κA whereas all edges have infinite capacity. Note that the capacity of each vertex can alternatively be represented by "splitting up" that vertex to two adjacent vertices of infinite capacity each and connecting them with an imaginary edge of capacity κA.

If the above network flow problem has a feasible solution, then there exists a pattern of renegotiations such that all banks would avoid default. In particular, the flow value on the edge connecting bank i to bank j is exactly the amount that one bank pledges to pay the other at $t = 2$, in exchange for a reverse transfer at $t = 1$. Also note that the assumption that vertex capacities are equal to κA guarantees that no bank can get a transfer at $t = 1$ which exceeds the value of its pledgeable assets at $t = 2$.

Given the above network flow problem, the max-flow min-cut theorem implies that a feasible flow exists as long as the value of the minimum cut in the network is larger than the total outgoing flow to the sinks, i.e., $m(\epsilon + v - a)$. Since the minimum cut would consists of at least one of the imaginary edges of capacity κA, the theorem implies that a feasible flow exists as long as $\kappa A > m(\epsilon + v - a)$, completing the proof.

Proof of part (b). Suppose that there exists a pattern of renegotiations which guarantees that no defaults occur. If all defaults are indeed avoided, any distressed bank i is paid back in full by its debtors, which means that i would have a total shortfall of $v + \epsilon - a$ on all its debt (to its senior creditors and other banks). Bank i would be able to renegotiate this debt only if it can pledge as much at $t = 2$; that is, only if $\kappa A > v + \epsilon - a$. This is, however, in contradiction with the assumption that $\epsilon > \min\{\epsilon^*, \kappa A + a - v\}$. Thus, at least one bank defaults. ∎

---

[36]For a careful treatment of network flow problems, see for example, Cormen, Leiserson, Rivest, and Stein (2001). For another application of this theory in economics, see Ambrus, Mobius, and Szeidl (2012).

References

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2012), "The network origins of aggregate fluctuations." *Econometrica*, 80, 1977–2016.

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2010), "Cascades in networks and aggregate volatility." NBER Working Paper 16516.

Allen, Franklin and Ana Babus (2009), "Networks in finance." In *The Network Challenge: Strategy, Profit, and Risk in an Interlinked World* (Paul R. Kleindorfer and Yoram (Jerry) Wind, eds.), 367–382, Wharton School Publishing.

Allen, Franklin, Ana Babus, and Elena Carletti (2012), "Asset commonality, debt maturity and systemic risk." *Journal of Financial Economics*, 104, 519–534.

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

Ambrus, Attila, Markus Mobius, and Adam Szeidl (2012), "Consumption risk-sharing in social networks." NBER Working Paper 15719.

Babus, Ana (2009), "The formation of financial networks." Discussion Paper 06-093, Tinbergen Institute.

Battiston, Stefano, Domenico Delli Gatti, Mauro Gallegati, Bruce Greenwald, and Joseph E. Stiglitz (2012), "Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk." *Journal of Economic Dynamics and Control*, 36, 1121–1141.

Berman, Abraham and Robert J. Plemmons (1979), *Nonnegative matrices in the mathematical sciences*. Academic Press, New York.

Billio, Monica, Mila Getmansky, Andrew W. Lo, and Loriana Pelizzon (2012), "Econometric measures of connectedness and systemic risk in the finance and insurance sectors." *Journal of Financial Economics*, 104, 535–559.

Blume, Larry, David Easley, Jon Kleinberg, Robert Kleinberg, and Éva Tardos (2011), "Network formation in the presence of contagious risk." *Proceedings of the 12th ACM Conference on Electronic Commerce*.

Brunnermeier, Markus K. and Martin Oehmke (2012), "Bubbles, financial crises, and systemic risk." In *Handbook of the Economics of Finance* (George M. Constantinides, Milton Harris, and René M. Stulz, eds.), volume 2, North Holland.

Brunnermeier, Markus K. and Lasse Heje Pedersen (2005), "Predatory trading." *The Journal of Finance*, 60, 1825–1863.

Caballero, Ricardo J. and Alp Simsek (2012), "Fire sales in a model of complexity." Working Paper.

Cabrales, Antonio, Piero Gottardi, and Fernando Vega-Redondo (2012), "Risk-sharing and contagion in networks." Working Paper.

Castiglionesi, Fabio, Fabio Feriozzi, and Guido Lorenzoni (2010), "Financial integration and liquidity crises." Working paper.

Cifuentes, Rodrigo, Gianluigi Ferrucci, and Hyun Song Shin (2005), "Liquidity risk and contagion." *Journal of the European Economic Association*, 3, 556–566.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein (2001), *Introduction to Algorithms*. MIT Press, Cambridge, MA.

"Counterparty Controversy" (2012). *The Economist*. http://www.economist.com/node/21554522.

Dasgupta, Amil (2004), "Financial contagion through capital connections: A model of the origin and spread of bank panics." *Journal of the European Economic Association*, 2, 1049–1084.

Diamond, Douglas W. and Philip H. Dybvig (1983), "Bank runs, deposit insurance, and liquidity." *Journal of Political Economy*, 91, 401–419.

Diamond, Peter A. (1982), "Aggregate demand management in search equilibrium." *Journal of Political Economy*, 90, 881–894.

Eboli, Mario (2012), "A flow network analysis of direct balance-sheet contagion in financial networks." Working paper.

Eisenberg, Larry and Thomas H. Noe (2001), "Systemic risk in financial systems." *Management Science*, 47, 236–249.

Elliot, Matthew, Benjamin Golub, and Matthew O. Jackson (2012), "Financial networks and contagion." Working Paper.

Elsinger, Helmut (2009), "Financial networks, cross holdings, and limited liability." Oesterreichische Nationalbank Working Paper 156.

Financial Crisis Inquiry Commission (2011), "The Financial Crisis Inquiry Report: Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States." U.S. Government Printing Office, Washington, DC.

Ford, Lester Randolph and Delbert Ray Fulkerson (1956), "Maximal flow through a network." *Canadian Journal of Mathematics*, 8, 399–404.

Freixas, Xavier, Bruno M. Parigi, and Jean-Charles Rochet (2000), "Systemic risk, interbank relations, and liquidity provision by the central bank." *Journal of Money, Credit and Banking*, 32, 611–638.

Gorton, Gary and Andrew Metrick (2011), "Securitized banking and the run on repo." *Journal of Financial Economics (forthcoming)*.

Gorton, Gary B. (2010), *Slapped by the Invisible Hand: The Panic of 2007*. Oxford University Press, USA.

Haldane, Andrew G. (2009), "Rethinking the financial network." Speech delivered at the Financial Student Association in Amsterdam.

Holmström, Bengt and Jean Tirole (1998), "Private and public supply of liquidity." *Journal of Political Economy*, 106, 1–40.

Jackson, Matthew O. and Asher Wolinsky (1996), "A strategic model of social and economic networks." *Journal of Economic Theory*, 71, 44 – 74.

Kiyotaki, Nobuhiro and John Moore (1997), "Credit chains." Working Paper.

Krishnamurthy, Arvind (2010), "Amplification mechanisms in liquidity crises." *American Economic Journal: Macroeconomics*, 2, 1–30.

Leitner, Yaron (2005), "Financial networks: Contagion, commitment, and private sector bailouts." *The Journal of Finance*, 60, 2925–2953.

Longstaff, Francis A. (2010), "The subprime credit crisis and contagion in financial markets." *Journal of Financial Economics*, 97, 436 – 450.

Lorenzoni, Guido (2008), "Inefficient credit booms." *The Review of Economic Studies*, 75, 809–833.

Plosser, Charles I. (2009), "Redesigning financial system regulation." A speech at the New York University Conference "Restoring Financial Stability: How to Repair a Failed System".

Rotemberg, Julio J. (2011), "Minimal settlement assets in economies with interconnected financial obligations." *Journal of Money, Credit and Banking*, 43, 81–108.

Shin, Hyun Song (2008), "Risk and liquidity in a system context." *Journal of Financial Intermediation*, 17, 315 – 329.

Shin, Hyun Song (2009), "Securitisation and financial stability." *The Economic Journal*, 119, 309–332.

Shleifer, Andrei and Robert W. Vishny (1992), "Liquidation values and debt capacity: A market equilibrium approach." *The Journal of Finance*, 47, 1343–1366.

Sorkin, Andrew Ross (2009), *Too Big to Fail: The Inside Story of How Wall Street and Washington Fought to Save the Financial System—and Themselves*. Viking, New York.

Stulz, René M. (2010), "Credit default swaps and the credit crisis." *Journal of Economic Perspectives*, 24, 73–92.

Vivier-Lirimont, Sébastian (2006), "Contagion in interbank debt networks." Working Paper.

Zawadowski, Adam (2011), "Interwoven lending, uncertainty, and liquidity hoarding." Working Paper.

Zawadowski, Adam (2012), "Entangled financial systems." Working Paper.