Systemic Risk in Endogenous Financial Networks*

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

This version: November 2014
First version: October 2014

Abstract

We provide a framework to study the formation of financial networks and investigate the interplay between the banks' lending incentives and the emergence of systemic risk. We show that under natural contracting assumptions, banks fail to internalize the implications of their lending decisions for the banks with whom they are not directly contracting, thus establishing the presence of a *financial network externality* in the process of network formation. We then illustrate how the presence of this externality can function as a channel for the emergence of systemic risk. 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 not spread their lending sufficiently among the set of potential borrowers, creating insufficiently connected financial networks that are excessively prone to contagious defaults. Finally, we show that banks' private incentives may lead to the formation of financial networks that are overly susceptible to systemic meltdowns with some small probability.

Keywords: systemic Risk, financial network, network formation, contagion.
JEL Classification: G01, D85.

*This paper is partially based on the previously unpublished results in the working paper of Acemoglu, Ozdaglar, and Tahbaz-Salehi (2013). We are grateful to Dean Corbae, Michael Gofman and Jennifer La'O for helpful feedback and suggestions. Acemoglu and Ozdaglar gratefully acknowledge financial support from the Army Research Office, Grant MURI W911NF-12-1-0509.

[†]Department of Economics, Massachusetts Institute of Technology.
[‡]Laboratory for Information and Decision Systems, Massachusetts Institute of Technology.
[§]Columbia Business School, Columbia University.

1 Introduction

Since the recent financial crisis, the view that interlinkages within financial markets can serve as a channel for propagation of shocks and lead to the emergence of systemic crises has gained traction among many commentators, economists, and regulators. For example, Janet Yellen (2013), at the time the Vice Chair of the Fed, noted that

> "Experience — most importantly, our recent financial crisis — as well as a growing body of academic research suggests that interconnections among financial intermediaries are not an unalloyed good. Complex interactions among market actors may serve to amplify existing market frictions, information asymmetries, or other externalities."

Not surprisingly, over the past few years, a large body of work has focused on the role of various types of interlinkages between financial institutions as mechanisms for the propagation and amplification of shocks. These studies, which for the most part model interbank interactions by the means of a network, analyze whether and how stress at a few institutions can spread to others via interbank linkages and result in systemic meltdowns.

Most of this literature takes the patterns of interbank interactions as exogenously given and provides a comparative analysis of the likelihood of systemic crises as a function of the structure of the economy's underlying financial network. This is despite the fact that financial institutions enter into contracts with one another voluntarily, consequently implying that financial interlinkages, and hence the structure of the resulting networks, are themselves equilibrium objects that are determined endogenously. This means that any network-based regulatory intervention that does not treat the network structure as part of the equilibrium is subject to the Lucas Critique (Lucas, 1976): the introduction of such policies may affect the banks' lending and borrowing incentives and hence, alter the underlying network of financial interlinkages. Thus, the proper assessment of any policy that relies on the structure of interbank linkages needs to be coupled with a theory of network formation that takes the lending and borrowing incentives of the banks explicitly into account.

In this paper, we provide one such framework to study the formation of financial networks and investigate the interplay between the banks' lending incentives and the emergence of systemic risk. Though stylized, our framework enables us to explicitly model the lending and borrowing decisions of the banks, analyze the terms of equilibrium contracts, and identify the potential inefficiencies that equilibrium financial networks exhibit.

We consider an economy consisting of n financial institutions (henceforth, banks for simplicity) that lasts for three periods. In the initial period, banks need to borrow from one another to invest in projects with (random) returns in the intermediate and final dates. Even though such bilateral lending relations enable the banks to realize the gains from trade, they also create a complex web of interbank obligations, thus exposing the banks to the risk of financial contagion.

The key endogeneity in our model thus involves the structure and the terms of bilateral interbank agreements. To capture this endogeneity concretely, 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 these covenants ensures that the extent of counterparty risk is reflected in the interbank contracts, in the sense that banks with a higher risk of default face higher interest rates in equilibrium. This feature distinguishes our model from the rest of the literature on the formation of financial networks in which prices and the allocation of surplus are exogenously determined.

As our first result, we show that thanks to the contingency covenants, all bilateral externalities are internalized, in the sense that each bank takes the impact of its actions on its *immediate* counterparties into account. To show this result in the most transparent manner, we focus on an architecture in which no other forms of network externalities are present. More specifically, we show that in an economy consisting of three banks, the fact that interbank interest rates adjust endogenously ensures that the extent of interbank lending in equilibrium coincides with the (constrained) efficiency benchmark.

We then argue that, in the presence of counterparty risk and the possibility of financial contagion, such an efficiency result may no longer hold. More specifically, we show that as long as banks cannot write extremely complicated contracts that are contingent on the intricate details of the financial network, the networks that are formed in equilibrium may not be socially efficient, thus implying that, in general, private and public incentives for forming financial connections may not coincide. This is despite the fact that interbank interest rates are adjusted endogenously to (partially) reflect the extent of counterparty risk. This observation highlight the presence of a novel form of externality, which we refer to as the financial network externality, in the process of network formation: 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.

As our next set of results, we illustrate the implications of this externality for the types of inefficiencies that may arise in the network formation process and the overall level of systemic risk. In particular, we first show that banks may "overlend" in equilibrium, as they do not internalize the fact that extending loans to their potential borrowers create channels over which idiosyncratic shocks can translate into systemic crises via financial contagion. This means that the social surplus can be increased if some banks refrain from lending, thus creating shorter credit chains that are more resilient to the risk of financial contagion.

Next, we show that the financial network externality can also manifest itself in the form of financial networks that are under-diversified, in the sense that banks may not spread their lending sufficiently among their potential counterparties. In particular, any given bank may find it beneficial to lend to fewer counterparties if such an under-diversified lending strategy reduces the bank's counterparty risk. Yet, at the same time, a more concentrated lending pattern leads to longer default cascades, thus increasing the overall level of systemic risk considerably. This result thus highlights that even if the total level of lending is efficient in equilibrium, the distribution of interbank obligations that endogenously arise may be inefficient from the social surplus perspective, exposing the system to excessively high levels of systemic risk.

As our final result, we show that banks' private incentives may lead to the formation of financial networks that are overly susceptible to systemic meltdowns. More specifically, we consider a highly interconnected financial network and show that if large shocks are rare enough, banks do not internalize the fact that their lending decisions may pave the way for transforming such rare, large shocks into systemic events in which a large number of banks default. This is despite the fact that systemic meltdowns are extremely costly from the social welfare perspective, no matter how rare they are.

Related Literature Our paper belongs to the growing literature that studies the interplay of financial networks and systemic risk. Dating back to the works of Allen and Gale (2000) and Freixas, Parigi, and Rochet (2000), this literature focuses on whether and how interbank interlinkages can function as a mechanism for the propagation and amplification of shocks. Some of the more recent studies, such as Amini, Cont, and Minca (forthcoming), Acemgolu, Ozdaglar, and Tahbaz-Salehi (2015), Elliott, Golub, and Jackson (2014) and Cabrales, Gottardi, and Vega-Redondo (2014), provide a comparative analysis of the extent of financial contagion as a function of the structure of the economy's underlying financial network. For instance, building on the model of Eisenberg and Noe (2001), Acemgolu et al. (2015) characterize the most and least fragile financial networks and show that highly interconnected financial networks exhibit a form of phase transition: even though such networks are highly stable at the face of small shocks, the presence of a large enough shock can lead to high levels of systemic risk.[1] For the most part, however, these papers take the structure of the financial network as exogenously given. In contrast, the main focus of our paper is to provide a framework for the endogenous formation of interbank linkages and study the implications of such endogeneity from a systemic perspective.

Our paper also belongs to the literature on strategic network formation, which dates back to the seminal works of Jackson and Wolinsky (1996) and Bala and Goyal (2000). This literature, which for the most part focuses on reduced-formed network formation games, studies how agents trade off the costs and benefits of creating links with one another and characterizes the set of networks that are formed in equilibrium. As a recent example, Erol and Vohra (2014) study a reduced-form network formation game that precedes a binary action game with local externalities and show that the structure of the equilibrium networks depends on whether the shocks to the system are believed to be correlated or independent from one another. In contrast to most of this literature, our framework explicitly models interbank contracts, thus enabling us to investigate the interplay between counterparty risk and the equilibrium interest rates.

More closely related to our work is a small subset of papers, such as Babus (2014), Zawadowski

[1]Some of the other papers in this literature include Gai, Haldane, and Kapadia (2011), Gofman (2011, 2014), Battiston, Delli Gatti, Gallegati, Greenwald, and Stiglitz (2012), Caballero and Simsek (2013), Cohen-Cole, Patacchini, and Zenou (2013), Georg (2013), Alvarez and Barlevy (2014), Denbee, Julliard, Li, and Yuan (2014), Glasserman and Young (2015) and Alter, Craig, and Raupach (2015).

(2013) and Farboodi (2014), which explicitly focus on the formation of financial networks. For example, Babus (2014) 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 that are highly resilient to the propagation of shocks. Zawadowski (2013), 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.

Our paper is most closely related to the recent work of Farboodi (2014), who studies a model of endogenous intermediation among debt financed banks. She shows that if only a subset of banks have access to a profitable investment opportunity, equilibrium financial networks have a core-periphery structure, in the sense that most banks enter into contracts with a few highly connected financial institutions. Furthermore, she shows that due to the presence of intermediation rents, equilibrium financial networks may not be efficient as social and private incentives are not necessarily aligned with one another. Despite their similarities, our work differs from Farboodi's in one significant way: even though both papers consider the endogenous formation of financial linkages, Farboodi takes the face value of the contracts and the allocation rule of intermediation rents as exogenously given. In contrast, all interbank interest rates in our model are determined endogenously. As we will show, this ensures that the extent of counterparty risk is reflected in the equilibrium interbank contracts.

In a more recent study, Di Maggio and Tahbaz-Salehi (2014) consider a networked model of secured lending, in which, similar to the current paper, the terms of the interbank contracts (that is, interest rates and haircuts) are endogenously determined. They show that by allowing the banks to write secured lending contracts, the equilibrium patterns of interbank lending are constrained efficient. In contrast to their work, we focus on the type of inefficiencies that may arise due to the presence of the financial network externality.

Finally, on the empirical side, Di Maggio, Kermani, and Song (2014) investigate how the network of relationships between dealers shapes their trading behavior and liquidity provision in the corporate bond market. They show that the market exhibits a clear core-periphery structure and that being a highly connected and systemically important dealer becomes more valuable during periods of high uncertainty.

Outline of the Paper The rest of the paper is organized as follows. Section 2 contains our model of financial institutions and presents the process of network formation. In Section 3, we describe our solution concept. We present our main results in Section 4, where we show that due to the presence of a form of financial network externality, equilibrium financial networks may be excessively prone to the risk of financial contagion. Section 5 concludes. All proofs are presented in the Appendix.

2 Model

2.1 Financial Institutions

Consider a single good economy, consisting of n risk-neutral banks indexed $\{1, 2, \ldots, n\}$ 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 periods: $t = 0, 1, 2$. At the initial date, each bank i borrows funds from other banks or the outside financiers to invest in a bulky project of size k that yields returns at the intermediate and final periods.

Each bank is endowed with k units of capital at $t = 0$, which it can lend to the rest of the banks in the economy. Alternatively, the bank can keep ("hoard") its excess capital as cash and obtain a rate of return that we normalize to 1. However, in analogy to the "coconut" model of Diamond (1982), we assume that the bank cannot use its funds to invest in its own project. Rather, it needs to borrow from the rest of the banks or the outside financiers to be able to initiate an investment.[2]

There are exogenous constraints on the extent to which banks can borrow from one another. Such restrictions may be due to liquidity or maturity mismatch across banks, asymmetric costs of peer monitoring, absence of long-term interbank relationships, or pairwise commitment problems. Formally, we assume that bank j can borrow at most k_{ij} units of capital from bank i; that is, $\ell_{ij} \leq k_{ij}$, where ℓ_{ij} denotes the amount that bank i lends to bank j. Throughout the paper, we assume that $\sum_{j \neq i, 0} k_{ji} \geq k$ for all i, thus guaranteeing that banks can always raise enough funds from other banks if they wish to do so.

As an alternative, 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 periods $t = 0$ and $t = 1$ (e.g., they have access to a linear risk-free technology with return r realized at $t = 1$).

2.2 Interbank Lending and Debt Contracts

Interbank lending takes place dynamically through debt contracts signed at $t = 0$ that have to be repaid at $t = 1$. At the beginning of period $t = 0$, 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

[2] 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 investment opportunities and the funds required for undertaking them 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 funds available 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 access to investment opportunities. In this alternative setup, the fact that i 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. Even though equivalent qualitatively, we rely on the variant of the model described in the main body of the paper to abstract from dynamic considerations.

lender specifies the interest rates at which it lends to the borrower as a function of the borrower's lending behavior. More formally, the timing of events over $t = 0$ 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 $\hat{R}_{ij} \in \{R_{ij}, \varnothing\}$.[3]

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 pairwise interbank interest rate at $\hat{R}_{ij}(\ell_{j1}, \ldots, \ell_{jn})$, thus implying that the face value of j's debt to i is equal to $y_{ij} = \ell_{ij}\hat{R}_{ij}(\ell_{j1}, \ldots, \ell_{jn})$.

The contracting game above has two notable features. First, the presence of contingency covenants allows the lenders to charge different interest rates depending on the risk-taking behavior of their borrowers. As we will show in Section 4, this assumption ensures that the extent of counterparty risk is (partially) reflected in the interbank contracts, in the sense that banks with higher risk of default face higher interest rates in equilibrium. This feature distinguishes our model from the rest of the literature on financial network formation, such as Farboodi (2014), in which prices and the allocation of surplus are exogenously determined.

Second, note that even though the interest rate offered by bank i to bank j is contingent on the lending behavior of j, it is independent of the terms at which j contracts with its other counterparties as well as the behavior of the rest of the banks in the economy. This assumption rules out the possibility that banks can write complicated contracts whose terms are contingent on the intricate details of the economy's underlying financial network structure.

Finally, we remark that 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.

[3]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}.

2.3 Investments and Debt Repayments

Once interbank lending is complete, bank i invests in a project which has a random short-term return of $z_i \in \{a, a-\epsilon\}$ at $t = 1$, where ϵ can be interpreted as the size of a negative shock to the project's return, a. If held to maturity, the project also has a fixed, non-pledgeable long-term return of A at $t = 2$. The bank can liquidate its project prematurely after the realization of the short-term returns at $t = 1$. However, it can only recover a fraction $\zeta < 1$ of the project's full value through liquidation. For simplicity, throughout the paper we assume that $\zeta = 0$.[4]

Once the bank undertakes the project, it must also meet an outside liability of magnitude $v > 0$ at $t = 1$, which is assumed to have seniority relative to the its liabilities to other banks and the outside financiers. These more senior commitments may be claims by the bank's retail depositors, wages due to its workers, taxes due to the government, or secured claims by non-bank financial institutions. Therefore, the total debts (liabilities) of bank i at $t = 1$ is equal to $y_i + v$ where $y_i = \sum_{j \neq i} y_{ji}$.[5]

Debt repayments at $t = 1$ follow closely the model of Acemgolu et al. (2015). In particular, if bank j is unable to meet its obligations in full, it defaults and has to liquidate its project prematurely where the proceeds are distributed among its creditors. We assume that all 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.

2.4 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.

Note that even though the restrictions on the amount that each bank can borrow from others, i.e., $\{k_{ij}\}$, are given exogenously, the face values of interbank debt contracts $\{y_{ij}\}$, and hence, the underlying structure of the financial network are endogenously determined in equilibrium.

3 Equilibrium Concept

Given that interbank lending decisions take place prior to debt repayments, a variant of subgame perfect equilibrium is the natural solution concept for our setup. We start by defining an equilib-

[4] This assumption, which ensures that liquidation of the project does not generate enough funds for the bank to meet its liabilities, can be relaxed without affecting our qualitative results.

[5] Note that y_i captures the obligations of bank i to the rest of the banks in the economy as well as any potential debt to the outside financiers.

rium concept for the debt repayment subgame at $t = 1$ while taking the structure of the financial network as given. In line with the models of Eisenberg and Noe (2001) and Acemgolu et al. (2015), we define our $t = 1$ solution concept as a collection of mutually consistent interbank payments. We subsequently define the full equilibrium of the game.

3.1 Payment Equilibrium

Consider a subgame in which the face values of interbank obligations are given by $\{y_{ij}\}$ and that bank j has kept an amount $c_j = k - \sum_{i \neq j} \ell_{ji}$ of its excess capital as cash.

Let x_{js} denote the repayment by bank s on its debt to bank j at $t = 1$. The total cash flow of bank j is thus equal to $h_j = c_j + z_j + \sum_{s \neq j} x_{js}$. If h_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, all of its creditors get paid equal to the face value of their claims; that is, $x_{ij} = y_{ij}$ for all $i \neq j$. If, on the other hand, $h_j < v + y_j$, bank j defaults and its creditors are repaid less than face value. In particular, when h_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 $h_j \in (v, v + y_j)$, the debt repayments by bank j to its junior creditors are 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

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

where $[\cdot]^+$ 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 $\zeta = 0$.

We can thus define the $t = 1$ equilibrium of the game as a collection of mutually consistent interbank payments:

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.

The following result, proved in Acemgolu et al. (2015), shows that, regardless of the patterns of interbank lending and the face values of the contracts, a payment equilibrium always exists and is unique for a generic set of parameter values.

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

3.2 Full Equilibrium

The concept of payment equilibrium defined above captures interbank repayments at $t = 1$. We now define the full equilibrium of the game which not only captures interbank repayments at $t = 1$, but also the banks' endogenous decisions of lending and borrowing at $t = 0$.

Definition 2. 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.[6]

(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, unlike the rest of the literature, 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 following section, ensures that the most obvious form of bilateral externalities are internalized, thus providing a useful framework for analyzing financial network externalities.

3.3 Social Surplus and Constrained Efficiency

Given the above framework, the (utilitarian) social surplus in the economy is naturally defined as the sum of the returns to all agents at the final period. More specifically,

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

[6] 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.

where $T_i \leq v$ is the repayment of 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. Therefore, if the outside financiers do not lend to any of the banks, then $\pi_0 = 0$.

The notion of social surplus defined above captures how interbank lending leads to a trade-off between investment efficiency and the extent of counterparty risk. Note that on the one hand, interbank lending increases social surplus, as each bank's excess capital is utilized for investment in other banks' projects with an expected net return of $r > 1$ to the lender[7] — a more efficient use of capital compared to hoarding it as cash at a rate of return of 1. On the other hand, however, an increase in interbank lending subjects the lenders to a higher level of counterparty risk and thus increases their likelihood of default. Since the bank's default leads to the premature liquidation of its project at cost A, an increase in the extent of interbank lending may simultaneously increase the total social cost due to inefficient liquidations.

Note that social surplus u does not *explicitly* depend on the interbank interest rates, as debt repayments are simply transfers between different banks. However, to the extent that the interest rates impact the likelihood of bank defaults, they would impact the total social surplus in the economy.

Throughout the paper, we use the above measure of social surplus to define a notion of constrained efficiency. According to this notion, the social planner controls the lending decisions of the banks, but not the interbank interest rates, which are determined by the equilibrium behavior of the outside financiers. Note that 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 the number of defaults. Such a framework, clearly, does not constitute a reasonable benchmark for comparison.

4 Formation of Financial Networks

In this section, we present our main results on how the endogenous formation of financial networks can lead to the emergence of systemic risk. More specifically, we show that even though interbank interest rates adjust endogenously in order to reflect the risk that each bank's decisions impose on its counterparties, equilibrium financial networks may be excessively prone to the risk of financial contagion.

4.1 Bilateral Efficiency

We start our analysis by arguing that thanks to the contingency covenants, all bilateral financial externalities are internalized. To this end, we focus on an extremely simple architecture in which no other forms of network externalities are present.

Consider an economy comprising of three banks labeled $\{1, 2, 3\}$, each endowed with k units of

[7] Note that since the outside financiers are always willing to lend at an (expected) rate of return of r, all other banks would also obtain the same expected return — in equilibrium — whenever they lend to one another.

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. Thus, if no bank relies on the outside financiers for funding, the three-chain financial network depicted in Figure 1 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. These parametric assumptions guarantee that, in the three-chain financial network, a negative shock to bank 1 leads to the failure all banks.

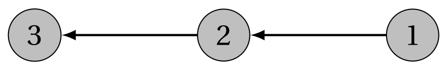


Figure 1. The three-chain financial network.

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; therefore implying that the lending decision of bank 2 (to bank 1) imposes an externality on bank 3. The following result, however, shows that thanks to the contingency covenants in the debt contracts, bank 2 fully internalizes this externality.

Proposition 2. *The three-chain financial network is part of an equilibrium if and only if it is (constrained) efficient.*

The above proposition thus establishes that each bank takes the effects of its actions on its immediate creditors into account. Put differently, it shows that all *bilateral* externalities are internalized, as each bank's creditors can offer contracts whose terms induce the borrower to take the "right" action.

To see the intuition underlying the above result more clearly, recall that the bilateral lending agreement between banks 1 and 2 increases the risk of contagion to bank 3. However, if, this increase in the counterparty risk is beyond the socially efficient level, bank 3 would be willing to offer bank 2 a sufficiently low interest rate provided that the latter refrains from making a loan to bank 1. This in turn ensures that bank 2 would lend to bank 1 only when bank 3 (as well as the social planner) would prefer this transaction to happen, thus restoring efficiency.

4.2 Financial Network Externality: Overlending

The stylized example of the three-chain financial network shows that, thanks to the covenants, borrowers internalize the externality that their actions impose on their immediate creditors. In the remainder of this section, we argue that this efficiency result may not hold if the economy's underlying the financial network has a more complicated structure. Rather, in the presence of counterparty risk

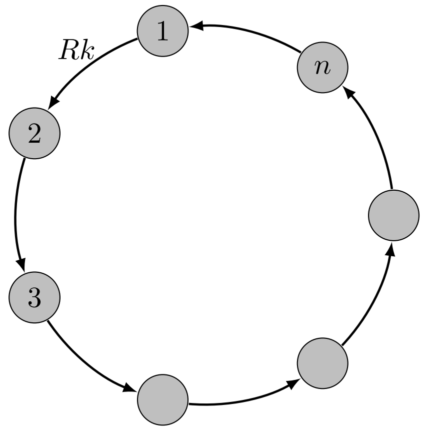


Figure 2. The ring financial network

and the possibility of financial contagion, private and public incentives for forming financial connections do not generally coincide, thus increasing the extent of systemic risk beyond the socially optimal level.

We start by showing that equilibrium financial networks may exhibit “overlending”, in the sense that banks lend to one another even though the social planner would have preferred that they hoarded cash to limit systemic risk.

To illustrate the nature of these inefficiencies in the most transparent fashion, we focus on a special configuration of interbank lending opportunities. Consider an economy in which 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 2 would form.

Even though stylized, focusing on the ring financial network has a few advantages. First, note that this structure is a natural extension of the three-chain financial network to an economy consisting of n banks, while at the same time ensuring that all banks take symmetric roles in the network. This symmetry in turn enables us to characterize the terms of equilibrium contracts in a closed-form fashion. Second, as our following discussion highlights, it captures the types of inefficiencies that are at the core of our results in the most transparent manner.

Proposition 3. *Suppose that one bank is hit uniformly at random with a shock $\epsilon < n(a-v)$. 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 (constrained) inefficient if $(r-1)k < \bar{\alpha} A$.

The ring financial network is thus part of an equilibrium if the interest rates that banks can charge when they lend to one another is large enough to justify the subsequent increase in the expected cost of a default, which is proportional to A.[8] On the other hand, part (b) of the above result shows that the

[8]Recall that the benefit of lending over hoarding cash is equal to $(r-1)k$.

ring financial network would be socially inefficient whenever the costs associated with the higher risk of financial contagion are so high that it is no longer justified for all banks to lend to their respective borrowers.[9]

The more important consequence of Proposition 3, however, 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 part of an equilibrium, even though it is inefficient from the point of view of the social planner.

The juxtaposition of Propositions 2 and 3 clarifies that the inefficiency identified above is not due to a simple bilateral externality. Rather, it is due to the presence of a more complicated externality, which we refer to as the *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. More specifically 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 $\tau > 2$, thus potentially impacting banks $i+2$, $i+3$ and so forth. Under such conditions, the social planner can increase social welfare by forcing one bank (say, bank $i+1$) to refrain from lending to its potential borrower (i.e., bank i), as such an action would shorten the average length of contagious defaults in the economy, decrease the likelihood of defaults of banks $i+2,\dots,i+\tau$, and as a result, reduce the overall level of systemic risk.

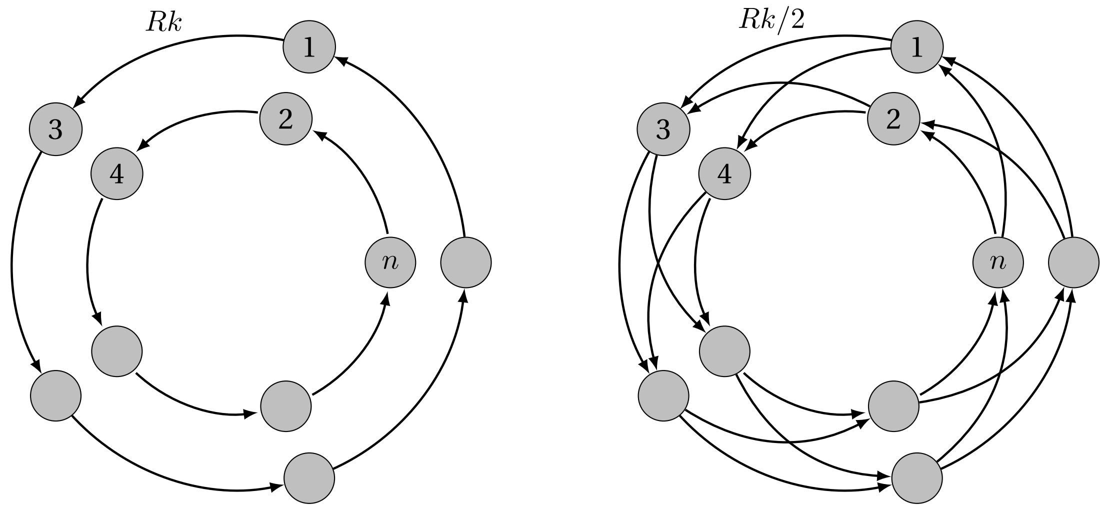
To summarize, the presence of the type of financial network externality identified above implies that financial stability is a public good that is under-provided in equilibrium. Furthermore, this inefficiency can manifest itself in the form of overlending in equilibrium: by lending to their respective borrowers, banks take risks that are deemed excessive from a social welfare perspective as such actions expose the banking system to longer chains of default.

4.3 Financial Network Externality: Under-Diversification

Proposition 3 shows that due to the presence of the financial network externality, banks may overlend in equilibrium relative to the (constrained) efficiency benchmark. In this subsection, we argue that the financial externality identified above may manifest itself in a different way. In particular, we show that equilibrium financial networks may be insufficiently dense, in the sense that — from the social planner's point of view — banks may not spread out their lending enough among all potential borrowers.

To illustrate the possibility of under-diverisfication, 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)

[9] The assumption that $\epsilon < n(a-v)$ is meant to guarantee that not all banks in the ring financial network default at the face of the shock. For more on this, see Acemgolu et al. (2015).



(a) The double-ring financial network (b) The interlinked rings financial network

Figure 3. The double-ring and the interlinked rings financial networks

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 3(b), would emerge. 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 3(a). As in the ring financial network studied in the previous subsection, the symmetric nature of these networks enables us to obtain a closed form expression for equilibrium interest rates and highlight the underlying insights of our model in the most transparent manner. We have the following result:

Proposition 4. *Suppose that one bank is hit uniformly at random with a shock $\epsilon < n(a-v)/2$. Then, there exists $\alpha > 0$ such that*

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

(b) The double-ring financial network is socially (constrained) inefficient.

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 3, the assumption that $\alpha A < (r-1)k$ guarantees that banks in the double-ring financial network find it optimal to lend to a single borrower. In particular, this inequality ensures that the expected returns on lending exceeds the expected cost of default due to the contagion of counterparty risk, thus implying that no bank would deviate by hoarding cash. Note that unlike the ring financial network in Subsection 4.2, each bank in the double-ring financial network can also deviate by following a diversified lending strategy. Yet, as part (a) shows, the bank does not find such a deviation to be profitable either, as lending to a diversified set of banks would expose the bank (and its immediate creditor) to a higher level of counterparty risk.

More importantly, however, Proposition 4 illustrates that even though no bank finds diversified lending profitable, such a deviation imposes a positive externality on the rest of the system. In particular, as part (b) shows, the double-ring financial network formed in equilibrium is inefficient from the social planner's perspective.

The intuition underlying this result relies on the observation that the extent of financial contagion is reduced when banks spread out their lending among more counterparties. Note that under such a diversified lending strategy, the burden of any potential losses is shared among more counterparties, hence, guaranteeing that the excess liquidity of the non-distressed banks can be more efficiently utilized in forestalling further defaults.[10] Consequently, by reducing the length of the failure cascades, such a diversified lending strategy benefits other banks further up the credit chain. Yet, bank i does not internalize this effect when deciding on its lending strategy.

To summarize, Proposition 4 shows that in addition to the possibility of overlending and excessive risk-taking, the presence of the financial network externality can also manifest itself in the form of socially inefficient levels of diversification, thus creating a new channel for the emergence of systemic risk.

4.4 Systemic Crises

We end this section by arguing that the financial network externality identified above can also lead to the emergence of financial networks that are overly susceptible to systemic meltdowns. More specifically, we show that even though extremely important from the social welfare perspective, banks do not internalize that their lending decisions may pave the way for transforming rare, large shocks into systemic crises in which a large number of banks default.

To capture how such systemic crises may arise endogenously in the most transparent manner, we focus on a particular configuration of interbank lending opportunities, where each bank i can borrow up to an amount $k_{ji} = k/(n-1)$ from any other bank $j \neq i$. Therefore, if no bank decides to hoard cash, the *complete financial network* depicted in Figure 4 would emerge. Before proceeding with our main discussion, we first present the following result, proved in Acemgolu et al. (2015):

Proposition 5. *Consider the complete financial network and let $\epsilon^* = n(a-v)$. Also suppose that a single bank is hit with a negative shock of size ϵ.*

(a) If $\epsilon < \epsilon^*$, the complete network has the minimal number of defaults, where only one bank fails.

(b) If $\epsilon > \epsilon^*$, the complete network has the maximal number of defaults, where all banks fail.

Part (a) of the above result thus shows that if the size of the negative shock is small enough, the complete network is the most stable financial network in which no bank (other than the originally distressed bank) defaults. The intuition underlying this result is that, in the complete network, the losses of the distressed bank are divided among as many creditors as possible, guaranteeing that the

[10] For more on this, see Acemgolu et al. (2015).

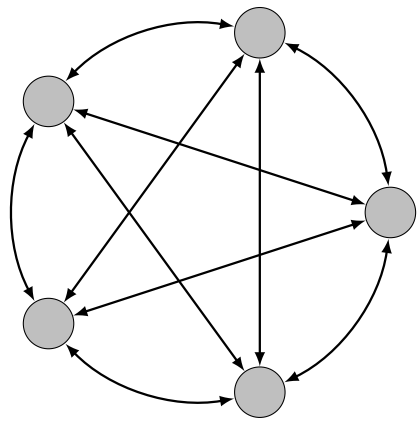


Figure 4. The complete financial network

excess liquidity in the financial system can fully absorb the transmitted losses. Part (b), however, shows that this logic does not generalize if the size of the shock passes the threshold ϵ^*. In particular, it shows that at the face of all such shocks, the complete network is the most fragile of all financial networks, where contagious cascades lead to the failure of all banks. Thus, taken together, the two statements of Proposition 5 highlight the "robust-yet-fragile" property of highly interconnected financial networks: such network structures are very resilient in response to a range of shocks. However, once we move outside this range, dense interconnections act as a channel through which shocks to a single financial institution transmit to the entire system, creating a vehicle for instability and systemic risk.

In the remainder of this section, we use Proposition 5 to show that banks may not internalize the likelihood of rare, systemic crises in which a large number of banks default. To capture the possibility of such crises, we 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. We have the following result:

Proposition 6. *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 (constrained) efficient.

(b) If $p > 0$ and $pA > \xi$, the complete financial network is socially (constrained) 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 5(a) that the complete financial network is extremely resilient in the presence of small shocks, as only the 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, however, is no longer true when there is a positive probability that a large shock is realized. As part (b) of Proposition 6 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 5(b). Recall that conditional on the realization of the large shock $\epsilon_h > \epsilon^*$, all banks in the complete financial network default. Thus, if the cost of default, A, is sufficiently large, the social planner would prefer a less interconnected network or even the empty configuration — in which banks withhold lending altogether — to the complete financial network. Such weakly connected or fully disconnected (empty) architectures ensure that the realization of a large shock at some distant corner of the network does not translate into a systemic crisis, affecting all banks.

Finally, part (c) of Proposition 6 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 are fairly unlikely, banks find lending more profitable than hoarding cash.

The key insight in Proposition 6 is obtained by comparing the conditions in its 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.[11] However, as soon as there is the slightest possibility of the realization of a large shock, public and private incentives for interbank lending start to diverge. 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 such a wedge in public and private incentives is the key reason behind the inefficiency of the complete financial network when $p > 0$.[12]

To summarize, Proposition 6 shows that in the presence of large, rare shocks, financial networks that are formed in equilibrium may be overly susceptible to contagious defaults, thus leading to the endogenous emergence of systemic risk.

5 Conclusions

In this paper, we provide a framework to study the formation of financial networks and investigate the interplay between the banks' lending incentives and the emergence of systemic risk. Unlike the rest of the literature on network formation, our model enables us to simultaneously endogenize the patterns of interbank lending as well as the face value of the banks' obligations to one another. More specifically, we assume that interbank interest rates adjust endogenously in order to reflect the risk that each bank's decisions impose on its immediate counterparties, while ruling out the possibility of writing complicated contracts that depend on the intricate details of the financial network's structure.

[11] Recall that when $p = 0$, the complete financial network fully absorbs the shock without any contagion.

[12] Statements (b) and (c) of Proposition 6 imply that the complete financial network can emerge 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 values of 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.

Our main results establish that equilibrium financial networks may be excessively prone to the risk of financial contagion. More specifically, we show that even though banks fully internalize the implications of their actions for 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. Our results thus highlight the presence of a financial network externality, which cannot be internalized via simple bilateral contractual relations. Furthermore, our results lend support to the view that financial stability is a "public good," that is likely to be under-provided in equilibrium.

The paper also investigates how the presence of the financial network externality can function as a channel for the emergence of systemic risk. In particular, it shows 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 not spread their lending sufficiently among the set of potential borrowers. Finally, we show that banks' private incentives may lead to the formation of financial networks that are overly susceptible to systemic meltdowns due to rare, large shocks.

A Appendix: Proofs

Throughout the proofs, 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.

Proof of Proposition 2

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

$$(r-1)k \geq pA. \tag{2}$$

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

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

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 (2) holds.

We next show that (2) 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.[13] 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

$$\begin{aligned} rk &= (1-p)R_{01}k + p(a - v - \epsilon + k) \quad (3) \\ rk &= (1-p)R_{02}k + p(2(a-v) - \epsilon + k) \quad (4) \\ rk &= R'_{02}k, \end{aligned}$$

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

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

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[14]

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

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 (2) does not hold, the 3-chain financial network cannot be part of an equilibrium.

The proof is complete once we show that (2) is also a sufficient condition for the 3-chain financial network to be part of an equilibrium. In particular, we show that if (2) holds, then the collection 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 (3) and (4), 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.

[14] 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}.

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 (2) 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. $\square$

Proof of Proposition 3

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 $i \in \{0, 1, \dots, n\}$ 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:

$$rk = \left(\frac{n-\tau}{n}\right) Rk + \frac{1}{n}\sum_{s=1}^{\tau}\Big[s(a-v) - \epsilon + Rk\Big] \tag{5}$$

$$rk = \left(\frac{n-1}{n}\right) R'k + \frac{1}{n}(a - v - \epsilon + k). \tag{6}$$

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

$$\pi_{\text{ring}} = \left(\frac{n-\tau}{n}\right) A + a - v - \frac{\epsilon}{n}, \tag{7}$$

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,

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

Thus, such an interest rate exists only if

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

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. $\square$

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,

$$\begin{aligned}\mathbb{E}[\tau_{\text{hoard}}] &= \frac{1}{n}\Big[(n - \tau - 1)\tau + \sum_{s=1}^{\tau} s + \tau'\Big]\\ &\leq \tau + 1 - \frac{\tau(\tau+1)}{2n},\end{aligned}$$

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

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

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

Proof of Proposition 4

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. Let $\tau = \lceil \epsilon/(a-v) \rceil - 1$ denote the number of 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,

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

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, as long as $\tau = \lceil \epsilon/(a-v) \rceil - 1$ is even, 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 $i \in \{0, 1, \ldots, n\}$ 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 (5) and (6).

Given that in the double-ring financial network each bank effectively belongs to a single credit chain, the results established in Proposition 3 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 (7) 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 $i-2\tau+4 \le j \le i$ 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 6

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 $0 \le s \le n-1$, 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.[15]

Lemma 1. *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 $R_s \le R_c$. 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

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

which is a contradiction. □

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

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, it is easy to see that 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

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

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 start by showing 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, as before, R_s denotes 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.

As a first step, note that given the posted contract $\mathbf{R}$, the complete network is a Nash equilibrium of the subsequent subgame. To see this, note that since 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 the banks. Furthermore, no bank has an incentive to deviate and lend less either, since by Lemma 1, such a deviation means that the bank would face a higher interest rate from its creditors, while getting a smaller rate of return on its own investment. Therefore, if all banks and the outside financiers post $\mathbf{R} = (R_0, \dots, R_{n-1})$, 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 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. $\square$

References

Acemgolu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2015), "Systemic risk and stability in financial networks." *American Economic Review*, 105.

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2013), "Systemic risk and stability in financial networks." NBER Working Paper No. w18727.

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

Alter, Adrian, Ben Craig, and Peter Raupach (2015), "Centrality-based capital allocations and bailout funds." *International Journal of Central Banking*.

Alvarez, Fernando and Gadi Barlevy (2014), "Mandatory disclosure and financial contagion." Federal Reserve Bank of Chicago Working Paper No. 2014-04.

Amini, Hamed, Rama Cont, and Andreea Minca (forthcoming), "Resilience to contagion in financial networks." *Mathematical Finance*.

Babus, Ana (2014), "The formation of financial networks." Discussion Paper 06-093, Tinbergen Institute.

Bala, Venkatesh and Sanjeev Goyal (2000), "A noncooperative model of network formation." *Econometrica*, 68, 1181–1229.

Battiston, Stefano, Domenico Delli Gatti, Mauro Gallegati, Bruce Greenwald, and Joseph E. Stiglitz (2012), "Default cascades: When does risk diversification increase stability?" *Journal of Financial Stability*, 8, 138–149.

Caballero, Ricardo J. and Alp Simsek (2013), "Fire sales in a model of complexity." *Journal of Finance*, 68, 2549–2587.

Cabrales, Antonio, Piero Gottardi, and Fernando Vega-Redondo (2014), "Risk-sharing and contagion in networks." Working Paper.

Cohen-Cole, Ethan, Eleonora Patacchini, and Yves Zenou (2013), "Systemic risk and network formation in the interbank market." Working Paper.

Denbee, Edward, Christian Julliard, Ye Li, and Kathy Yuan (2014), "Network risk and key players: a structural analysis of interbank liquidity." FMG Discussion Paper DP734.

Di Maggio, Marco, Amir Kermani, and Zhaogang Song (2014), "The value of trading relationships in turbulent times." Working paper.

Di Maggio, Marco and Alireza Tahbaz-Salehi (2014), "Financial intermediation networks." Columbia Business School Working Paper No. 14-40.

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

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

Elliott, Matthew, Benjamin Golub, and Matthew O. Jackson (2014), "Financial networks and contagion." *American Economic Review*, 104, 3115–3153.

Erol, Selman and Rakesh Vohra (2014), "Network formation and systemic risk." PIER Working Paper No. 14-029.

Farboodi, Maryam (2014), "Intermediation and voluntary exposure to counterparty risk." Working Paper.

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.

Gai, Prasanna, Andrew Haldane, and Sujit Kapadia (2011), "Complexity, concentration and contagion." *Journal of Monetary Economics*, 58, 453–470.

Georg, Co-Pierre (2013), "The effect of the interbank network structure on contagion and common shocks." *Journal of Banking & Finance*, 37, 2216–2228.

Glasserman, Paul and H. Peyton Young (2015), "How likely is contagion in financial networks?" *Journal of Banking & Finance*, 50, 383–399.

Gofman, Michael (2011), "A network-based analysis of over-the-counter markets." Working Paper.

Gofman, Michael (2014), "Efficiency and stability of a financial architecture with too-interconnected-to-fail institutions." Working Paper.

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

Lucas, Robert E. (1976), "Econometric policy evaluation: A critique." In *Carnegie-Rochester Conference Series on Public Policy* (Karl Brunner and Alan Meltzer, eds.), volume 1, 19–46, Elsevier.

Yellen, Janet L. (2013), "Interconnectedness and systemic risk: Lessons from the financial crisis and policy implications." Remarks at the American Economic Association/American Finance Association joint luncheon, San Diego, California.
http://www.federalreserve.gov/newsevents/speech/Yellen20130104a.pdf.

Zawadowski, Adam (2013), "Entangled financial systems." *The Review of Financial Studies*, 26, 1291–1323.