What Makes Financial Networks Special? Distorted Investment Incentives, Regulation, and Systemic Risk Measurement.

Matthew O. Jackson and Agathe Pernoud*

Draft: July 2019

Abstract

In a model of financial networks with both debt and equity interdependencies, we show that financial organizations have incentives to: choose excessively risky portfolios; overly correlate their portfolios with those of their counterparties; and under-diversify in terms of choosing too few counterparties with whom to share risk. We also provide a measure of financial centrality in terms of how a given organization's portfolio affects the values and defaults of other organizations. Additionally, we characterize optimal regulation in terms of the use of reserve requirements versus bailouts, and fully characterize the minimum bailouts needed to ensure systemic solvency.

JEL Classification Codes: D85, F15, F34, F36, F65, G15, G32, G33, G38

Keywords: Financial Networks, Markets, Systemic Risk, Financial Crisis, Correlated Portfolios, Networks, Banks, Default Risk, Financial Interdependencies

1 Introduction

World trade has grown from just under 20 percent of world GDP at the end of the Second World War to over 60 percent.[1] This unprecedented growth in trade has had many benefits from various forms of gains from trade, economies of scope and scale, more efficient investment, and has been accompanied by a comparable growth in the international financial network. For instance, the amount of investment around the world coming from foreign

*Department of Economics, Stanford University, Stanford, California 94305-6072 USA. Jackson is also an external faculty member of the Santa Fe Institute. Email: jacksonm@stanford.edu and agathep@stanford.edu. We gratefully acknowledge financial support under NSF grant SES-1629446. We thank Marco Bardoscia, Celso Brunetti, Ozan Candogan, Matthew Elliott, Gerardo Ferarra, Co-Pierre Georg, David Hirshleifer, Arjun Mahalingam, Carlos Ramirez, Tarik Roukny, and Alireza Tahbaz-Salehi, for helpful conversations and comments.

[1]Imports plus exports over GDP. Detailed data can be found for 1870-1949: Klasing and Milionis (2014), 1950-1959: Penn World Trade Tables Version 8.1, 1960-2015: World Bank World Development Indicators.

sources went from 26 trillion dollars in 2000 to over 132 trillion dollars in 2016, which today represents more than a third of the total level of world investments.[2] In addition, the financial sector is characterized by strong interdependencies, with capital circulating from financial firm to financial firm. Using administrative data from the US Federal Reserve Bank, Duarte and Jones (2017) estimate that 23% of the assets of bank holding companies come from within the US financial system, as well as 48% of their liabilities - *almost half*.[3,4] Along with the enormous benefits that have accompanied the growing and increasingly inter-connected world economy, have come stronger conduits of shocks and risks of widespread contagion. These are not idle concerns, as we witnessed in 2008 when exposure to a problematic mortgage market led to key insolvencies in the US and elsewhere, and to a broad financial crisis and prolonged recession.[5]

Financial markets differ from textbook efficient markets on several dimensions, and are important to understand since they are fundamental to all businesses and sectors of the economy. Financial markets are ripe with externalities, often subtle but with wide-ranging consequences. At a most basic level, the risk that a counterparty defaults has consequences in a world with some missing markets in which not all risks can be fully hedged.[6,7] Defaults involve substantial inefficiencies and costs, which result from fire sales, early termination of contracts, government bailouts, and legal costs, among others; much of which are born by parties others than those who are responsible for the original default. Even without full cascades, these costs are substantial. For example, estimates of bankruptcy recovery rates are in the 56-57 percent range.[8] Defaults can then lead to potential disasters if left to cascade. These introduce further externalities, since if one organization has poor judgment in its investments, poorly managed business practices, or simply unusually bad luck, this ends up affecting the values of its partners, and their partners, in discontinuous ways. Costs of avoiding cascades and other externalities can be much less if addressed before they begin. However, this involves having a detailed view of the network of financial inter-dependencies,

[2]See IMF publication: Susan Lund and Phlippe Harle "Global Finance Resets" in *Finance and Development*, Dec 2017, volume 54 number 4, pp 1-4.

[3]The large difference reflects the fact that many other types of financial organizations that are not BHCs (e.g., Real Estate Investment Trusts, Insurance Companies, and various sorts of investment funds, etc.) have accounts of cash, money markets, and other deposits held at BHCs.

[4]Also, between 1980 and 2018 there was an enormous consolidation. The number of banks has dropped to a third of what it was, and at the same time banks are managing more than eight times as much in terms of total assets. See Jackson and Pernoud (2019) for more background.

[5]For narratives of the crisis see the US Congressional *Financial Crisis Inquiry Report* of January 2011, as well as Glasserman and Young (2016) and Jackson (2019).

[6]As a poignant example, there are even risks that the organizations offering insurance and hedges default. For instance, a key failure in the financial crisis in 2008 was that AIG was unable to deliver the insurance it had sold on many derivative contracts. Its inability to even meet margin calls on that insurance, and subsequent insolvency, forced a large government intervention.

[7]In terms of relations to textbook general equilibrium and efficient markets, not only are markets incomplete, but they also involve substantial discontinuities that can even preclude existence, as we discuss in Section A.1.

[8]See Branch (2002); Acharya, Bharath and Srinivasan (2007), as well as Davydenko, Strebulaev and Zhao (2012); James (1991).

an understanding of the consequences of those interdependencies, as well as of the incentives that different parties have in choosing their investments and counterparties, and thus their position in the network. Although there is a growing literature on financial networks and their consequences,[9] the incentives of financial organizations to choose their investments are not well-understood.

In this paper, we examine the distortions of financial organizations' incentives to invest due to being embedded in a network. We do this in the context of a new model of financial networks that involves two different forms of contracts: debt and equity. This generalizes existing models used to understand systemic risk that are either built with debt-like interdependencies – Eisenberg and Noe (2001), Gai and Kapadia (2010)), and Csóka and Herings (2018) – or with equity-based interdependencies – Elliott, Golub and Jackson (2014)). Our model is more in the tradition of those that have been used in the accounting and asset valuation literatures (Suzuki (2002); Fischer (2014)) to derive equity prices when firms also issue debt. Our model generalizes those models to include multiple types of contracts, while also allowing for bankruptcies and resulting discontinuities due to bankruptcy costs, since those are central to the inefficiencies and externalities in financial networks. We highlight that accounting for both debt and equity contracts is important, as they lead to very different incentives for investment as well as different probabilities of cascades, for the same initial conditions.[10]

We then examine three basic choices of banks and other sorts of financial organizations: which investments they make, how many counterparties they transact with, and the extent to which they choose to correlate their portfolios with those of their counterparties. We show how externalities result in inefficiencies in all of these choices, and discuss the network consequences. Our first result states that, under general conditions on the network structure, banks choose to take on too much risk as compared to what is socially efficient. This comes from the fact that banks do not account for the negative externalities their default imposes on the overall financial system when trading-off the benefits and costs of a risky investment. We use this analysis to examine optimal regulation of portfolio choices in the

[9]A partial list of references is Rochet and Tirole (1996), Kiyotaki and Moore (1997), Allen and Gale (2000), Eisenberg and Noe (2001), Upper and Worms (2004), Cifuentes, Ferrucci and Shin (2005), Leitner (2005), Allen and Babus (2009), Lorenz, Battiston, Schweitzer (2009), Gai and Kapadia (2010), Wagner (2010), Billio et al. (2012), Elliott, Golub and Jackson (2014), Demange (2016), Diebold and Yilmaz (2014), Dette, Pauls, and Rockmore (2011), Gai, Haldane, and Kapadia (2011), Greenwood, Landier, and Thesmar (2012), Ibragimov, Jaffee and Walden (2011), Upper (2011), Allen, Babus and Carletti (2012), Cohen-Cole, Patacchini and Zenou (2012), Gouriéroux, Héam and Monfort (2012), Alvarez and Barlevy (2015), Glasserman and Young (2015), Acemoglu, Ozdaglar and Tahbaz-Salehi (2015), Gofman (2017), Babus (2016), Bardoscia (2017), Cabrales, Gottardi, and Vega-Redondo (2017), Erol and Vohra (2018), Kanik (2018).

[10]Although debt and equity capture only some of the main forms of interdependencies that one observes in practice, they provide a lens into many others, as many swaps and derivatives involve essentially either fixed payments or payments that depend on the realization of the value of some investment of one of the parties. Also, things like syndicated loans and other joint investments have features that are similar to equity. Debt and equity capture the two primary situations: contracts that are fixed in payments and those in which payments depend on value realizations.

form of reserve requirements, bailouts, or laissez-faire. Larger potential gains from the risky investment favor laissez-faire, and the use bailouts when network contagion becomes likely, while smaller gains favor reserve requirements. Interestingly, more debt actually favors laissez-faire and bailouts, while greater equity in the network favors reserve requirements, since the opportunity cost of requiring reserves scales with the amount of debt that an organization holds. On the extensive margin, we show that financial organizations also choose to have fewer counterparties with which to cross-insure that would be socially optimal, again since their owners do not bear the full brunt of their potential bankruptcy costs. In addition, we show that banks have strong incentives to correlate their investments with those of their counterparties, which can be seen as another form of excessive risk-taking. This happens for a slightly different reason, and we refer to it as 'risk-stacking', as it helps them align their solvencies with situations in which they can enjoy better payments from their counterparties, and being insolvent when they expect lower potential payments from their counterparties.

We also discuss how the model can be used to measure systemic risk from a potential shock or default; and we identify the minimum costs and specific interventions a government needs to bail out an insolvent network. Our results fully characterize the conditions needed for solvency under both the best and worst equilibria, as there can exist multiple equilibria. A corollary provides conditions for uniqueness of equilibria, and our results also point out the importance of canceling out cycles of debts among banks.

Here we emphasize that the method that we argue is necessary to properly manage systemic risk is conceptually very intuitive, but that it requires detailed information about the full network: having partial balance sheets from firms – without the precise listing of all counterparties – is insufficient to properly manage the network. We illustrate that point via an example, showing how it can be impossible to approximate systemic risk without detailed network information.

The literature looking at how financial organizations that are embedded in a network choose their portfolios is scarce. A notable exception is an (independent) study by Elliott, Georg and Hazell (2018), who also highlight an incentive of banks to choose partners that have portfolios similar to their own. They examine which network of inter-bank equity claims and correlation structure of investments arise endogenously in equilibrium when banks act under limited liability. Their main result is that banks have an incentive to partner with counterparties whose portfolios are positively correlated with their own in order to shift losses from shareholders to creditors. Using data on the German banking system, Elliott, Georg and Hazell (2018) also provide strong empirical evidence that banks lend more to other banks with portfolios similar to their own. Thus, such correlations are observed. Similar incentives arise in our model for partly different reasons, and persist even if banks do not have limited liability. In our model, because of financial interdependencies, banks' values depend positively on each other, which induces complementarities in their returns to investments: a high return for a bank is weakly more valuable if its partners have high returns as well, pushing them to correlate their portfolios.[11,12]

Though further away from what we study in this paper, the recent literature on the inefficiency of network formation in financial settings is also worth noting. The equilibrium network often has a core-periphery structure,[13] and is generally socially inefficient, either because it induces excessive systemic risk[14] or too much market power of core organizations.

2 A Model of Financial Interdependencies

Here, we define a model of financial interdependencies that allows us to examine the choices of financial organizations in the network. In the model we include both debt and equity since there are important distinctions in the incentives and systemic risk that they generate on the network. Insolvencies are induced by inability to pay debts, but not by equity. Thus, the relative combination of debt and equity in a bank's portfolio has important implications for incentives, and so it is important to allow for that distinction in the model.

The distinction between debt and equity is not just a theoretical consideration, since both types of securities are needed to capture the balance sheets of some of the most prominent and important types of financial organizations. For example, banks' balance sheets involve substantial portions of deposits, loans, CDOs (collateralized debt obligations), and other sorts of debt-like instruments. In contrast, venture capital firms and many other sorts of investment funds typically hold equity and are either held privately or issue equity. Such funds, and other forms of shadow-banking, are increasingly important as a source of funding for businesses, especially in the tech sector and other growing parts of the economy. Furthermore, some large investment banks are hybrids that involve substantial portions of both types of exposures. Understanding the different incentives that these different forms of organizations have, and the externalities that are present, is thus relevant.

[11] Acharya and Yorulmazer (2007) highlight a different channel that can drive banks to correlate their investments: if it is ex-post optimal for the regulator to bail-out banks when many of them fail at once, but to let them rescue each other when only few of them are insolvent, then banks have an incentive to herd in order to capture the bailout subsidies. A similar intuition arises in Arya and Glover (2001) in a principal-agent model, in which agents may want to coordinate on the bad action—e.d. low effort—if this means a higher probability of being bailed-out by the principal. In those papers the inefficiency stems from limited commitment of a principal or regulator, rather than from network externalities. Finally Erol (2019) shows that the anticipation of bailouts leads banks to form a highly concentrated network, which entails greater output volatility and systemic risk.

[12] Duffie and Wang (2016) study whether bilateral bargaining over the terms of contracts between banks achieves efficiency. Assuming away general cross-network externalities—and hence the possibility of default cascades—they propose a bargaining protocol that leads to socially efficient equilibrium contracts between banks. They however highlight that small changes to the proposed protocol may lead to additional inefficient equilibria.

[13] See Soramäki, Bech, Arnold, Glass and Beyeler (2007); Bech and Atalay (2010); Erol (2019); Blasques, Bräuning and Van Lelyveld (2018). There are a variety of reasons to have a core-periphery structure as there are advantages to having a concentration in intermediaries, which can then better manage their inventory and match buyers with sellers (e.g., see Craig and Von Peter (2014); Babus and Hu (2017); Farboodi (2017); Wang (2017)).

[14] For instance, see Elliott, Golub and Jackson (2014).

Many forms of contracts, including some swaps, can be approximated as some combination of debt and/or equity. Nonetheless, there are obviously more complex contracts that can also be built into such a model. We also describe a general form of the model for such more complex contracts in an appendix, but the main insights regarding incentives are most crisply analyzed with debt and equity, which captures the main tradeoffs and still remains piecewise-linear and hence tractable.

2.1 Banks, Shadow Banks, CCPs, and other Financial Organizations

Consider a set $N = \{0, 1, \ldots, n\}$ of organizations involved in the network. We treat $\{1, \ldots, n\}$ as the financial organizations, or "banks" for simplicity in terminology. These should be interpreted as a broad variety of financial organizations, including banks, venture capital funds, broker-dealers, CCPs (central counterparties), insurance companies, and many other sorts of shadow banks and institutions that have substantial financial exposures on both sides of their balance sheets. These are organizations that can issue as well as hold debt, buy and sell equity, and make other investments. The broad applicability of the model allows it to be used to assess risks of the evolving roles of CCPs and the large shadow banking sector, and not just traditional banks.

We lump all other actors into 0 as these are entities that either hold debt and equity in the financial organizations (for instance private investors and depositors), or borrow from or raise money from the financial organizations (for instance, most private and public companies). Their balance sheets may be of interest as well, as the defaults on mortgages or other loans could be important triggers of a financial crisis. The important part about 0 is that, although these may be the initial trigger and/or the ultimate bearers of the costs of a financial crisis, they are not organizations that are the dominoes, becoming insolvent and defaulting on payments as a result of defaults on their assets.[15]

Each organization i has a value V_i, which is the total value of the equity: the value of all investments including those in other organizations, net of all debts owed. This value is shared among private shareholders and institutional shareholders. We now describe these values in greater detail.

[15]There is a spectrum that involves a lot of gray area. For instance, Harvard University invests tens of billions of dollars, including making large loans. At the same time it borrows money and has issued debt of more than five billion dollars. It is far from being a bank, but still has incoming and outgoing debt and other obligations. This is true of many large businesses, some that come closer to resembling banks than others. Also, some companies' solvency could be affected by other bankruptcies and bring down counterparties, especially if they are key players in a supply chain. In that case, those companies would be included in the main $1, \ldots, n$, while companies that are usually just borrowers or just lenders and not potential dominoes are the actors in 0. It is not so important for us to try to draw an arbitrary line through this grey area to make the points that we do with our model. Nonetheless, this is something that a regulator does have to take a stand on when trying to assess systemic risk, but will often be dictated by jurisdictional rules.

2.2 Primitive Assets, Organizations, and Cross-Holdings

Bank portfolios are composed of both investments in primitive assets outside the network and financial contracts within the network. For our purposes the details of investments in primitive assets are not important: suppose they involve some initial investment of capital and then pay off some cash flows over time, often randomly. We call these primitive investment opportunities *assets* – $M = \{1, \dots m\}$ – and denote by p_k the present value (or market price) of asset $k \in M$. The values of organizations are ultimately based on their investments in these assets. Let $q_{ik} \geq 0$ be the quantity invested in asset k by organization i, and $\mathbf{Q}$ the matrix whose (i, k)-th entry is equal to q_{ik}. (Analogous notation is used for all matrices.) The total value of i's direct investments in primitive assets is thus $\sum_k q_{ik}p_k$, or $q_i \cdot p$.

The book or equity value V_i of an organization i equals the value of organization i's primitive assets plus the debts it is owed minus those it owes plus the value of its claims on other organizations. For the sake of transparency and tractability in what follows, we restrict the set of financial contracts to debt and equity. In the appendix (Section A) we discuss how valuations work with fully general contracts across banks, and how existence depends on a monotonicity of those contracts in underlying primitive-asset investments.

If bank i owns an equity share in bank j, it is represented by $S_{ij}V_j$ for some $S_{ij} \in (0, 1)$. A debt contract with a current (net present) value of D_{ij} corresponds to a payment of D_{ij} as long as bank j is solvent, while it will look like an equity share if j becomes insolvent. A call option looks like a value of $D_{ij} = 0$ until V_j exceeds a certain value, and then looks like a claim S_{ij} above that level. In a world with debt and equity, then the value of organization i's investments in other organizations is then $\sum_j D_{ij} + S_{ij}V_j$ and its total debt liability is $\sum_j D_{ji}$. These together with the primitive investments determine the equity value, V_i, which is then owned by private shareholders and other financial organizations through their equity shares in i.

Let $\mathbf{D}$ the matrix of debt claims and $\mathbf{S}$ the matrix of equity claims, where D_{ij} is what organization j owes to i and S_{ij} is i's equity claim on j. Let $S_{ii} = 0$ for all i such that a bank cannot have an equity claim on itself. Organizations are either privately owned $\sum_i S_{ij} = 0$, or public $\sum_i S_{ij} = 1$. A financial network is then a tuple $(N, \mathbf{D}, \mathbf{S})$.

Let

$$D_i^A = \sum_j D_{ij} \quad \text{and} \quad D_i^L = \sum_j D_{ji}$$

denote the total amount of debt owed to i and owed by i, respectively. The former is then i's debt assets, and the later its liabilities.

2.3 Values in a Network of Debt and Equity

Let V_j^+ denote $\max[V_j, 0]$. Given that equity has limited liability, then the value to i of its equity holding in j is $S_{ij}V_j^+$.

The book or equity value V_i of an organization i can then be written as:

$$\begin{aligned} V_i &= \sum_k q_{ik} p_k + \sum_j (D_{ij} - D_{ji}) + \sum_j S_{ij} V_j^+ \\ &= \sum_k q_{ik} p_k + D_i^A - D_i^L + \sum_j S_{ij} V_j^+ \end{aligned} \tag{1}$$

Equation (1) can be written in matrix notation as

$$\mathbf{V} = \mathbf{Qp} + \mathbf{D}^A - \mathbf{D}^L + \mathbf{SV}^+. \tag{2}$$

To ensure a (unique) solution to (2), it is sufficient that there exists at least one private organization (e.g., at least one private citizen who owns some share of some organizations), and that every public organization has some indirect private ownership—i.e. there exists a directed path in equity from every public bank to some private investor. In the case in which all organizations are solvent, so that $\mathbf{V}^+ = \mathbf{V}$, solving (2)[16] then yields

$$\mathbf{V} = (\mathbf{I} - \mathbf{S})^{-1} \left[\mathbf{Qp} + \mathbf{D}^A - \mathbf{D}^L \right]. \tag{3}$$

We handle the case with insolvencies and bankruptcies below.

The special case in which $\mathbf{S} = \mathbf{0}$ corresponds to the model of Eisenberg and Noe (2001); Gai and Kapadia (2010) and the special case in which $\mathbf{D} = \mathbf{0}$ corresponds to Elliott et al. (2014).

Written this way, the book or equity value of a publicly held organization coincides with its total market value. Indeed as argued by both Brioschi, Buzzacchi, and Colombo (1989) and Fedenia, Hodder, and Triantis (1994), the ultimate (non-inflated) value of an organization to the economy – what we call the "market" value – is well-captured by the equity value of that organization that is held by its *outside* investors – or the *final* shareholders who are private entities that have not issued shares in themselves. This value captures the flow of real assets that accrues to final investors of that organization. This is exactly what is characterized by the above values since summing them up (again, for the case of nonnegative

[16] To see that $(I - S)$ is invertible, and to ensure that all of the V's are bounded, it is sufficient that when we examine the directed network defined by positive S_{ij}'s, every node in the network is path connected to some private node - so j that has no public equity. Without this condition, the Vs are indeterminate. Intuitively, without this condition, there are no real owners of some companies - there would be a cycle in which every firm's value is dependent on the other values in the cycle and they all are fully owned in the cycle, and then the values are no longer tied down by the fundamentals.

values) gives

$$\begin{aligned}
\sum_i V_i &= \sum_i \sum_k q_{ik} p_k + \sum_i D_i^A - \sum_i D_i^L + \sum_i \sum_j S_{ij} V_j \\
&= \sum_i \sum_k q_{ik} p_k + \sum_{j \text{ public}} V_j \\
\implies \sum_{i \text{ private}} V_i &= \sum_i \sum_k q_{ik} p_k
\end{aligned}$$

The total value to all private investors equals the total value of primitive investments.

2.4 Discontinuities in Values and Failure Costs

We now introduce bank defaults and their associated costs.

If the value of an organization i's assets falls below the value of its liabilities, then i is said to *fail* and incurs failure costs $\beta_i(\mathbf{V}, \mathbf{p})$. These costs could depend on the degree to which i and others are insolvent as well as the value of its various direct investments. In the case of debt and equity, an organization's liabilities are its debt obligations, and its assets include primitive investments and equity it holds in other organizations, combined with the value of debt it is owed by others.[17]

With the possibility of bankruptcy, a debt owed to i by organization j, D_{ij} depends on the value of V_j and thus its solution depends ultimately on the full vector $\mathbf{V}$. Thus, in these cases, we write $D_{ij}(\mathbf{V})$ to make these interdependencies explicit.

There are two regimes. If organization j remains solvent, it can repay its creditors in full, and then for all i

$$D_{ij}(\mathbf{V}) = D_{ij}.$$

If instead organization j defaults, then debt holders become the residual claimants in case of insolvency, and

$$D_{ij}(\mathbf{V}) = \frac{D_{ij}}{\sum_h D_{hj}} \max\left(\sum_k q_{jk} p_k + D_j^A(\mathbf{V}) + \sum_h S_{jh} V_h^+ - \beta_j(\mathbf{V}, \mathbf{p}), 0 \right).$$

A simple example of bankruptcy costs corresponds to the case where $\beta_j(\mathbf{V}, \mathbf{p})$ equals a share of the value of the bank's assets, such that as a bank goes into bankruptcy, it only recovers some fraction of its assets (e.g., due to a markdown on a firesale of its assets); but it could also include fixed amounts of legal and other costs. If the bank is unable to salvage any of its assets, then $D_{ij}(\mathbf{V}) = 0$ whenever j defaults.

In any case, equity holders of organization j do not receive any payment when j is

[17]The model extends to allow for other sorts of contracts and thresholds for defaults, by simply having a different rule for when default occurs, for instance in the case of more complicated liabilities.

insolvent:

$$S_{ij}V_j^+ = 0.$$

The bankruptcy costs incurred by organization i are then

$$b_i(\mathbf{V}, \mathbf{p}) = \begin{cases} 0 & \text{if } \sum_k q_{ik}p_k + \sum_j S_{ij}V_j^+ + D_i^A(\mathbf{V}) \geq D_i^L \\ \beta_i(\mathbf{V}, \mathbf{p}) & \text{if } \sum_k q_{ik}p_k + \sum_j S_{ij}V_j^+ + D_i^A(\mathbf{V}) < D_i^L. \end{cases} \tag{4}$$

Note that we have carefully written bankruptcy costs $b_i(\mathbf{V}, \mathbf{p})$ as a function of $\sum_k q_{ik}p_k + \sum_j S_{ij}V_j^+ + D_i^A(\mathbf{V}) - D_i^L$ rather than as a function of V_i directly. This avoids having bankruptcies driven solely by anticipating bankruptcy costs, even when an organization has more than enough assets, even cash on hand, to cover its liabilities.[18] Our formulation allows for coordination issues, bank runs, and other sorts of multiplicities in equilibria that are of practical interest, while avoiding more trivial self-fulfilling bankruptcies that would just be modeling curiosities.

The valuations in (3) have analogs when we include these discontinuities in value due to failures and bankruptcy costs. The discontinuous drops impose costs directly on organizations' balance sheets. and so the book value of organization i becomes:

$$V_i = \sum_k q_{ik}p_k + \sum_j S_{ij}V_j^+ + D_j^A(\mathbf{V}) - D_i^L - b_i(\mathbf{V}, \mathbf{p}),$$

where $b_i(\mathbf{V}, \mathbf{p})$ is defined by (4). This leads to a new version of (3):

$$\mathbf{V} = (\mathbf{I} - \mathbf{S}(\mathbf{V}))^{-1} \left(\left[\mathbf{Qp} + \mathbf{D}^A(\mathbf{V}) - \mathbf{D}^L \right] - \mathbf{b}(\mathbf{V}, \mathbf{p}) \right), \tag{5}$$

where $\mathbf{S}(\mathbf{V})$ reflects the fact that $S_{ij}(\mathbf{V}) = 0$ whenever j defaults, and is S_{ij} otherwise.

It can be that some solutions for the values are negative - as the bankruptcy costs could exceed the assets of the company. We can think of these as real costs, for instance a decaying and polluting plant which is abandoned, or some court costs, etc., which in some cases are absorbed by a government or imposed on the public.

2.5 Existence and Multiplicity of Organization Values

Under the assumptions mentioned above, and under the assumption that the bankruptcy costs β are nonincreasing in $(\mathbf{p}, \mathbf{V})$ (so that bankruptcy costs are weakly lower when organizations have greater values), there always exists a solution to equation (5). There can exist

[18] Such a self-fulfilling bankruptcy would go beyond a bank run, since it would not be due to the organization not having enough cash on hand to pay its debts, but instead due to a fixed point issue that if we presume all the cash is eaten up by paying bankruptcy costs, then indeed the organization can become bankrupt for no other reason. This self-fulfilling problem posed by bankruptcy costs is of less interest to us here, as it is not so much a network issue—it is not a bank-run problem, nor a problem of coordinating payments with some cycle of other banks,— and seems of less practical interest.

multiple solutions to the valuation equation (multiple vectors $\mathbf{V}$ satisfying (5)) in the presence of discontinuities, and in fact, the set of equilibrium values forms a complete lattice.[19] We use the term "equilibrium" to refer to a fixed point satisfying equation (5) to keep with the literature, but note that the term "equilibrium" is also used for situations in which we endogenize the portfolio and partner choices.

As highlighted by Elliott, Golub and Jackson (2014), there are two sources of multiplicity. The first one is due to self-fulfilling bank runs (see classic models such as Diamond and Dybvig (1983)): there can be an equilibrium in which some bank i is solvent and another one in which it defaults even when keeping everything else constant. The second source of multiplicity comes from bank interdependencies: there can exist an equilibrium in which a subset of banks is solvent and another in which they all default. This corresponds to self-fulfilling default cascades.

Since the equilibria form a complete lattice, there exists a "best" equilibrium, as well as a "worst" equilibrium in which the set of defaulting organizations is minimal (Elliott, Golub and Jackson (2014)). The following algorithm finds the best equilibrium. Assuming that every bank is solvent, compute the right-hand side of (5). If any value is negative, the associated bank defaults and the values are computed again accordingly. Iterating this process yields the best equilibrium. The worst equilibrium can be found using a similar algorithm, but initially assuming that all organizations default and then seeing which ones are solvent and iterating. We note that this worst equilibrium also corresponds to a situation in which no bank makes any partial payments until it has received all of its payments in.[20] Despite the fact that bank values depend on the details of how they make payments, as well as the rules and timing in bankruptcy proceedings, our results do not.

In this paper we account for the multiplicity of equilibria, and in some cases distinguish between results for the best and worst equlibria. In particular, studying the worst equilibrium for values after a first failure may be relevant given how financial markets are subjects to runs and freezes: the regulator may want to consider the worst that could happen under such circumstances. Since many of the results apply to both the best and worst equilibria (as well as intermediate selections), as long as one is consistent in using the same equilibrium throughout the analysis,[21] we are only explicit about which equilibrium applies when it becomes necessary.

[19] This can be seen by an application of Tarski's fixed point theorem, since organizations' values depend monotonically on each other.

[20] That situation is discussed by Bardoscia, Ferrara, Vause and Yoganayagam (2019).

[21] We rule out that a bank believes that the selection of equilibrium changes as a function of its decisions – for instance that the best equilibrium applies for some decisions and the worst for others, which could alter the incentives arbitrarily.

3 Distorted Incentives: Externalities and Inefficiencies in Portfolio Choices

The previous section focused on characterizing values of banks, taking as given their portfolio and the structure of the financial network. We now model banks' investment decisions, and the potential inefficiencies that arise from network externalities.

3.1 Overly Risky Investment: The Intensive Margin

Because of interdependencies between organizations in the financial network, a bank's investment decision not only affects its value but also those of others. This relates to the standard agency problem between the manager of a firm and its shareholders highlighted in Jensen and Meckling (1976), in which there are differing objectives of the manager and shareholders.[22]

In this section, we study how financial contracts between organizations distort investment incentives, because of associated externalities in insolvencies and bankruptcy costs. The distortions in incentives are easy to understand, but are important to document because of their extremity and also because this analysis provides the base for an analysis of regulation and bailouts, as well as how incentives depend on whether interdependencies are based on debt or equity.

Let us begin by examining the incentive problems of a single bank that takes as given investments made by other financial organizations. Without loss of generality, the bank has a unit of capital to invest either in a risk-free asset with net return r or in some portfolio that pays a random p_i. One can think of this as a standard two-fund separation setting, with the main decision of the investor being how much risk to take.

We take the bank owners to be risk neutral and choose the portfolio that maximizes their expected returns from their investments. This allows us to abstract away from results linked to risk tolerance and to analyze solely systemic and structural externalities. We comment below on how the results extend to the case of risk-aversion. We also presume that the decision-makers in an organization are maximizing its profits and not the profits of other financial organizations.

Suppose the bank's outstanding debt D_i^L is low enough, such that it could be paid back entirely were the bank to only invest in the safe asset: $D_i^L \leq (1+r)$. Insolvency happens if the value of the portfolio falls below the bank's liability D_i^L, in which case a cost of $b \leq D_i^L$ is incurred. Under limited liability, the shareholders get a payoff of zero in case of insolvency, and the bankruptcy cost b is born by whomever is holding debt, or a government or other enterprise that steps in.

[22]For more on this point, see Admati and Hellwig (2013). For some analysis in a different network setting see Brusco and Castiglionesi (2007), and for a more general discussion about agency problems of excessive risk taking in the presence of externalities see Hirshleifer and Teoh (2009).

The bank solves

$$\max_{q_i \in [0,1]} (1-s)\mathbb{E}\left[\left(q_i p_i + (1-q_i)(1+r) + \sum_{j \neq i} S_{ij} V_j(\mathbf{p}, q_i)^+ + D_i^A(\mathbf{p}, q_i) - D_i^L\right)^+\right].$$

Here we do not allow for short sales (of either asset), which limits $q_i \in [0,1]$. As will be clear below, the analysis extends to short sales: the bank would choose to short the risk-free asset and leverage its investment in the risky asset.

We allow $V_j(\mathbf{p}, q_i)$ and $D_i^A(\mathbf{p}, q_i)$ to depend on the vector of portfolio values $\mathbf{p}$, and also on q_i. We consider i's best reply, so the investment decisions of other agents are taken as given and built into $V_j(\mathbf{p}, q_i)$ and $D_i^A(\mathbf{p}, q_i)$. As we will see, i's best response will not depend on those decisions.

We say that i is "*at least as dependent upon its own portfolio as others*," if for any $\mathbf{p}$ at which a slight change from q_i to q_i' causes some j to become insolvent, then i must also be insolvent at q_i'.

PROPOSITION **1.** *Consider a setting in which the vector of portfolio values* $\mathbf{p}$ *has a bounded and atomless distribution, and consider the portfolio choice of some firm* i *such that* $\mathbb{E}[p_i] > 1 + r$, $\mathbb{E}[p_i] > D_i^L$, *and* i *is at least as dependent upon its own portfolio as others. Then the bank will invest entirely in the risky portfolio.*

Proposition 1 does not extend to some examples in which some other bank or organization j is more sensitive to i's portfolio than i is. For instance, suppose that j has large debts to i and that j has equity in i - so, for example, they did a swap of equity in i for debt in j. Suppose also that j is only solvent if i's return is above $1 + r$, while i has no debt and is always solvent. If p_i has a large probability of falling below $1+r$, but only offers a small gain in expected return, then i is better off investing mostly in the risk-free asset. So, incentives to take (excessively) risky positions can be mitigated in some settings. However, this is a sort of extreme example in which there *must* be a nontrivial chance of driving a counterparty in whom i has a large stake into bankruptcy *without* having i become bankrupt.

3.1.1 Inefficiency and Risk Aversion

Fully investing in the risky portfolio is often socially inefficient, since a bank's decision also affects the rest of the financial network. First, a bank does not account for default/bankruptcy costs when deciding its investment: the above maximization decision is independent of b. Indeed, under limited liability, the bank shareholders only consider returns earned when solvent and completely disregard what happens under insolvency. Second, a bank's investment decision impacts others through financial contracts and cross-holdings. In particular, if i defaults it will not honor its debt liabilities and its creditors may be driven to insolvency, causing bankruptcy costs to add up. Because of these, a planner would often prefer less risky investments.

The intuition behind this is straightforward. The the bank has incentives to maximize

$$\mathbb{E}[V_i|V_i > 0]\Pr[V_i > 0],$$

while the full impact on society is

$$\mathbb{E}[V_i|V_i > 0]\Pr[V_i > 0] + \mathbb{E}[V_i - b|V_i < 0]\Pr[V_i < 0] - [\text{Contagion Cost}]\Pr[V_i < 0],$$

where the second and third terms in the latter expression are negative and generally become more negative in the amount of risk taken in the investment, q_i.

This makes it clear that, although the results may be attenuated, they will extend to the case of risk averse investors – there are still negative externalities on others that are not taken into account by those choosing the investment.

As mentioned above, this sort of incentive problem is reminiscent of other settings with externalities, but now we explore more deeply some of its implications, including optimal regulation, and incentives to correlate investments in a financial network.

3.1.2 Optimal Regulation: Bailouts versus Deposit Requirements

To develop further intuition, we consider a case of proprietary assets that just take on two possible payoffs. Each bank has one unit of capital that it can either invest in its risky asset or in a safe one. Bank i's risky asset yields return R with probability θ and zero otherwise, while the safe asset pays the risk-free rate $1+r$. Let $\theta R > 1+r$, as otherwise the risky asset is dominated by the risk-free asset.

There is a critical level of investment in the risky asset $\bar{q}$ under which the bank remains solvent irrespective of the realization of the risky asset's return. This threshold solves

$$(1-\bar{q})(1+r) = d \quad \text{or} \quad \bar{q} = 1 - \frac{d}{1+r}.$$

Thus, the bank's owners payoffs are

$$(1-s)\theta(qR + (1-q)(1+r) - d)$$

if $q > \bar{q}$, and

$$(1-s)\left[q\theta R + (1-q)(1+r) - d\right]$$

if $q \leq \bar{q}$.

Given that $\theta R > (1+r)$, it follows directly that the bank's optimization problem has corner solution of $q = 1$, which is clearly above $\bar{q}$. The optimal portfolio is such that the bank becomes insolvent with probability $1-\theta$.

Next, let us examine the overall total value maximization problem that a social planner would solve. We consider two cases.

Reserve Requirements First, suppose that the planner only considers regulation that limits the amount invested in the risky portfolio, i.e., a reserve requirement. The bankruptcy cost b is incurred by someone even if the corporation has limited liability, so it will be whomever those costs are owed to or has to step in (e.g., government, courts, other creditors, etc.).

If $q \leq \bar{q}$ then total payoffs are

$$\theta(qR + (1-q)(1+r)) + (1-\theta)(1-q)(1+r).$$

If $q > \bar{q}$ then the total society payoffs are

$$\theta(qR + (1-q)(1+r)) + (1-\theta)((1-q)(1+r) - b).$$

Given the risk premium $\theta R - (1+r) > 0$, the two possible maximizers are $q = 1$ or $q = \bar{q}$. Indeed there are two possible scenarios: either it is optimal for the social planner to prevent default with $q^* \leq \bar{q}$, or it is not. There is inefficient over-investment by the bank when the solution to the social problem is $\bar{q}$, that is when

$$\bar{q}\theta R + (1-\bar{q})(1+r) > \theta R - (1-\theta)b.$$

This happens when

$$R < b\frac{(1-\theta)}{\theta(1-\bar{q})} + \frac{1+r}{\theta} \iff \theta R < (1+r)\left(1 + \frac{b}{d}(1-\theta)\right).$$

Thus, there is inefficiency - or a need for regulation of the portfolio/reserves - when:

$$\theta R < (1+r)\left(1 + \frac{b}{d}(1-\theta)\right). \tag{6}$$

The bank over-invests in the risky asset when the risk premium is not high enough to compensate losses due to default. Otherwise, the expected gain in overall investment in the risky asset outweighs the potential bankruptcy costs, and it is socially optimal to fully invest. Intuitively, over-investment is more likely to happen for higher bankruptcy costs b since the bank overlooks these costs when choosing its portfolio. The inefficient region is also decreasing in R and θ, and increasing in r: the social planner is more likely to find it optimal to allow full investment in the risky asset for higher values of the risky asset's expected return.

More surprisingly, the inefficient region is *decreasing* in the level of outstanding debt d and vanishes for $d \geq \frac{(1+r)(1-\theta)b}{\theta R - (1+r)}$. That is, the social planner is more likely to accept the possibility of default, and allow the bank to invest fully in the risky asset when there is greater debt. The intuition behind this is that as debt increases, a greater investment in the safe asset is required to avoid default, which is increasingly costly in terms of expected returns to investment.

Bailouts Next, consider a situation in which a government can offer a bailout, meaning it can intervene and pay a bank's total liability d when there is an insolvency. This then avoids the bankruptcy cost. There is a loss of d to the government, but this ends up transferred to someone else, so that does not count as a net loss. If such bailouts can be done at no cost, then this will be the efficient solution. The more relevant case is one in which intervention actually involves some costs, denoted $c > 0$. These are various administrative costs, and costs of taking over a failing organization, which can be substantial in their own right. Effectively, in terms of the total societal cost of insolvency, this substitutes the cost c for the cost b. We consider the case in which $c < b$, as otherwise bailouts are not improving (in this example, with no contagion).

The end result is that if

$$\theta R < (1 + r)\left(1 + \frac{c}{d}(1 - \theta)\right),$$

then bailouts are inefficient and the bank over-invests, and so the optimal policy is to regulate that the bank keep at least $1 - \bar{q}$ in risk-free assets. In contrast, if

$$\theta R > (1 + r)\left(1 + \frac{c}{d}(1 - \theta)\right),$$

then it is efficient to let the bank invest fully in the risky asset and to bail them out in the case that they become insolvent.[23]

We now give an example of investment decisions in a network of banks, which thus amplifies the potential cost of insolvency via default cascades.

An Example with Cascades

Consider a network composed of debt contracts, in which banks have one counterparty each. Label banks so that bank $i + 1$ has a debt liability to bank i. Figure 1 depicts the corresponding network when there are $n = 2$ banks for simplicity, but the following analysis extends to any larger n. Node 0 is a depositor who has a demand deposit in Bank 1 in the amount D_{01}, and Bank 1 has lent an amount D_{12} to Bank 2. Each bank has a unit of capital for investment, and the additional capital that the banks have beyond what they have in net liabilities (e.g., $1 - D_{12}$ for Bank 2) are the investments of the owners of the banks.

Just as in Section 3.1.2, each bank can invest either in a risk-free asset or in its proprietary investment opportunity that yields a return $p_i = R > 0$ with probability θ and nothing

[23] Bailouts may have side effects that decrease their usefulness, but those are beyond the scope of our model. For instance, if there is an agency problem and managers prefer to keep their jobs and are not incentivized to choose the portfolio that would be optimal from the shareholders' perspective, then providing a bailout might induce them to invest in the risky asset when without a bailout they would not do so. Such incentive misalignments add another layer to the potential for inefficiency.

otherwise. In this simple network, Proposition 1 applies and all banks want to fully invest in their risky asset, irrespective of how their returns correlate across banks.

Consider a case in which Bank 1 is a net creditor: $D_{12} > D_{01}$, and such that banks could ensure paying back their debt by fully investing in the safe asset: $D_{i,i+1} < 1 + r < R$. Given that banks fully invest in the risky asset, Figure 1 shows which banks are solvent (blue) and insolvent (magenta) for each possible realization of (p_1, p_2).

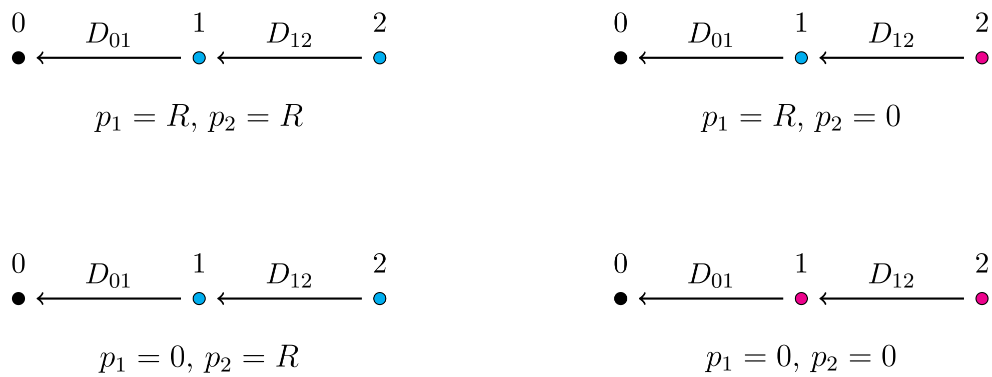


Figure 1: Solvent (blue) and insolvent (magenta) banks for each realization of asset returns (p_1, p_2) when each bank fully invests in its risky asset, and $D_{12} > D_{01}$.

The societal value of such financial network equals the expected total returns to investment minus bankruptcy costs—debt repayments are simply transfers from one agent to another, and hence do not enter efficiency calculations. Importantly, this value depends on how correlated banks' investment opportunities are.

To see this, suppose that they are perfectly correlated with probability α, and independent otherwise. Then the value to society under equilibrium risky investments is

$$2\theta R - b(1-\theta)[2-(1-\alpha)\theta]$$

which is decreasing in α. The first term captures expected returns, since no high return R is ever lost to an insolvent bank. The expected number of defaults $(1-\theta)[2-(1-\alpha)\theta]$ increases in the probability of perfect correlation α since correlation reduces the gains from a diversified network portfolio.

To highlight the inefficiency of investing fully in the risky asset in this example, suppose that instead bank 2 is obligated to have a portfolio sufficiently safe to guarantee its own solvency—i.e. it cannot invest more than $\bar{q}_2 \equiv 1 - \frac{D_{12}}{1+r}$ in its risky asset. Then bank 2 never defaults and, given this, bank 1 never defaults as well. Note that here it is sufficient to only regulate banks that are net debtors, while regulating net creditors only reduces the gains from investment without affecting expected bankruptcy costs. The societal value of such regulated network is

$$\theta R + \theta R \left[1 - \frac{D_{12}}{1+r}\right] + D_{12},$$

where the first term is the expected return of bank 1's portfolio, and the two last terms that of 2's portfolio. Hence there is excessive risk-taking as soon as

$$\theta R + \theta R \left[1 - \frac{D_{12}}{1+r}\right] + D_{12} > 2\theta R - b(1-\theta)[2-(1-\alpha)\theta],$$

which can be rewritten as

$$\theta R < (1+r)\left(1 + \frac{b}{D_{12}}(1-\theta)[2-(1-\alpha)\theta]\right). \tag{7}$$

It is interesting to compare this to the inefficiency calculation with just one bank, which appears in (6). The right hand side has the extra term $[2-(1-\alpha)\theta]$. We note two things about this factor. First, the factor 2 captures the fact that there are now 2 banks and bankruptcy costs at stake instead of one, and a potential contagion from Bank 2 to Bank 1 (the difference between the two lower situations in Figure 1), and more generally would have a factor n in the term if there are n banks involved. Second, this term is increasing in the correlation, which makes it more likely that Bank 2 pushes Bank 1 into insolvency - as if Bank 2 is solvent when Bank 1 gets no return on its portfolio, Bank 1 will still be solvent. The correlation increases the chances that this occurs.

This example points out a main issue, both in how Bank 2's investments now have additional inefficiencies, and that correlation can exacerbate that inefficiency. We show in Section 3.2 that, in fact, banks have incentives to correlate their investments.

3.1.3 Investment Incentives Under Debt vs. Equity

Banks have lower default thresholds when their liabilities are in the form of equity rather than debt. In the above analysis, including Proposition 1, if none of the liabilities were in the form of debt, then there would never be any bankruptcy costs or inefficiencies. Keeping incentives constant, equity is then more efficient since it reduces the probability of default, and hence the expected bankruptcy costs that the system could have to pay. The contrast between debt and equity also imply that the systemic risk roles of organizations like banks whose balance sheets have large amounts of debt on both sides, differ from that of venture capital and other funds whose balance sheets are almost exclusively equity-like.

There are, however, reasons for using debt: to pay workers who may have their own fixed bills to pay, to account for the risk aversion of investors, and to handle short term loans and demand deposits, etc. For instance a risk-averse investor would prefer to have a fixed payment than a random one, for the same expected value.[24]

An Example with Countervailing Incentives

There are situations in which debt can offer a countervailing incentive that leads to less risky investments, if there is a special sort of feedback effect in the network. We illustrate this effect in the following example. Essentially, it has to be that a low portfolio return would not directly cause the bank to become insolvent, but would instead lead one of its equity-holding counterparties to become insolvent, which would then lead to a loss of a payment back to the bank itself. The following example illustrates this intuition, and highlights the fact that incentives depend on the solutions for values; e.g., whether banks expect the worst- or best-case equilibrium for values.

Consider two banks $i = 1, 2$. Bank 1 owes $\underline{d}$ to bank 2, while 2 owes $\overline{d}$ to 1 with $\underline{d} < \overline{d}$. Furthermore bank 2 owns some equity share s of bank 1. Suppose bank 2 can only invest in the safe asset, but that this investment is not enough to cover its debt obligation to bank 1. Bank 1 on the contrary can decide which fraction of its portfolio to invest in a risky asset that yields R with probability θ.

First note that in the best equilibrium—i.e., the bank values with as few defaults as possible—default never occurs. Indeed even if bank 1 has a portfolio return of zero, it can still be solvent if 2 pays back its debt. Similarly 2 is always solvent if 1 pays back its debt. Hence if bank 1 expects the best-case values to arise, then its optimal portfolio is to only invest in the risky asset and set $q^* = 1$.

However in the worst equilibrium both banks default if $(1+r)(1-q) < \underline{d}$, which discontinuously decreases bank 1's expected value. There are then two candidates for the optimal portfolio: $q^* = 1$ and $q^* = 1 - \frac{\underline{d}}{1+r}$. The former is optimal if and only if

$$\theta R \left[1 - \frac{\underline{d}}{1+r}\right] + (1+r)\frac{\underline{d}}{1+r} - \underline{d} + \overline{d} \geq \theta \left[R - \underline{d} + \overline{d}\right]$$

or simplifying,

$$\overline{d}(1-\theta) \geq \theta \underline{d} \left[\frac{R-(1+r)}{1+r}\right],$$

that is, if the gain from having its debt repaid in all states more than compensates the loss in risk premium from the safer portfolio.

This example illustrates some of the nuances of financial interdependencies. Which mix of debt and equity is best for incentives depends on multiple features. Both debt and equity generally incentivize banks to fully invest in the risky portfolio. For them to choose safer investments, it is necessary to have a mix of debt and equity that allows for discontinuous

[24]Once bankruptcy costs are involved, this would no longer be the case, since a fixed payment would also be variable and have a lower expected value, although also a lower variance. The optimal contract in the face of a risk-averse investor and bankruptcy costs, could turn out to be a hybrid of debt and equity, or all one or the other, depending on the bankruptcy costs and variability of the investment portfolio and level of risk-aversion.

indirect effects of one's return on its own value through the network, and this can also depend on which equilibrium for values is expected.[25]

3.2 Correlated Investments: Popcorn *and* Dominoes

The metaphor of "popcorn or dominoes" was made by Eddie Lazear, the chairman of the council of economic advisors under Bush during the financial crisis. The question was whether there really was any issue of potential contagion and "dominoes", or whether much of the crisis was instead simply due to all banks "boiling in the same hot oil" - i.e. all having extensive exposure to an under-performing mortgage market. The answer is that both were true. Banks had very correlated portfolios and all had dangerously low values in their investments at the same time, and hence most were either barely solvent, or even insolvent. Nonetheless, they also had large exposures to each others' debts, as well as to derivatives from AIG, who could not even manage margin payments, as well as securities issued by Fannie Mae and Freddie Mac, which were both insolvent. This made it clear that a large cascade would occur without intervention.[26]

This highlights the fact that correlation of investments across banks matters for financial contagion: the fact that many organizations held directly or indirectly similar subprime mortgages made the whole system substantially more fragile. This can also happen through things like syndicated loans and other partnerships, swaps, and general incentives to hold the same assets. Correlation of investments affects systemic risk in two ways: it makes the network more prone to contagion conditional on a first failure, but can also change the probability of a first failure. Indeed, under correlated investments, as discussed above, if a bank gets a low or negative return on its investment then it is likely that its counterparties are in a similar situation. Equity claim and/or debt claims then may also pay weakly less, which increases the probability that the bank defaults. If it does default, its counterparties are also more likely to become insolvent since they also face insufficiently high enough asset returns to absorb the shock.

The point that banks can *prefer* to be partnered with other organizations that have similar portfolios was first made by Elliott, Georg and Hazell (2018), in which banks choose to lend to organizations with similar exposures so as to correlate their defaults. This also appears in a particularly robust and simple form in our model, and so we illustrate it now. We discuss how this result relates to theirs below.

We model a bank's choice of portfolio as a choice of states in which to get high returns, and allow banks to arbitrarily correlate their returns by choosing how much the sets of states in which they get high returns overlap - the flip side of choosing to whom to lend.

The idea behind our result is simple, which is why it turns out to be so robust: financial

[25] This is not necessary for these indirect effects, which can also occur in the best equilibrium, but this example also points out how the equilibrium that applies matters.

[26] For discussion of this see Jackson (2019), as well as the extensive analysis and data in the *Financial Crisis Inquiry Report*, commissioned by an act of the US congress.

organizations prefer to be solvent when their counterparties earn highest returns in order to enjoy part of those returns, and prefer to be insolvent when their counterparties are insolvent since then there are then lower returns coming in indirectly.

The importance of these results is that, regardless of whether one believes there is any serious contagion across banks, their incentives to correlate investments lead to coordinated failures and large losses for the economy at the same time.

3.2.1 An Example with Two Banks

Consider two banks – each have debt d to some outside investors and a (net) share of s in each other. Suppose there exists two independently distributed risky assets yielding a return of R_1 and R_2, respectively, with same probability θ, where $R_i > d$ for each i. Each bank can choose in which portfolio of these two assets it wants to invest.

To understand the equilibrium, we analyze two cases. The first is such that a bank that gets 0 becomes insolvent, but that a bank that earns a positive return stays solvent. The second is such that both banks are solvent if either gets a positive return, and then both are insolvent only when they both get 0 returns. There is a third case in which if either gets a 0 then both become insolvent, and it is straightforward that they prefer to correlate their portfolios then, so we do not do the incentive calculations for that case.

We start with the first case in which a bank that gets 0 becomes insolvent, but that a bank that earns a positive return stays solvent. Suppose, for now, that bank 1 is fully invested in asset 1. Then bank 2 wants to choose the same portfolio if

$$\theta \frac{(1+s)\left[R_1 - d\right]}{(1-s^2)} > \theta^2 \frac{\left[R_2 + sR_1 - (1+s)d\right]}{(1-s^2)} + \theta(1-\theta)\left[R_2 - d\right].$$

This simplifies to

$$R_1[1 + s(1-\theta)] > R_2[1 - s^2(1-\theta)] + s(1+s)(1-\theta)d.$$

This is always true when $R_1 \geq R_2$,[27] and holds even if $R_1 < R_2$, if d is sufficiently small.

Next, consider the second case in which both banks are solvent if either gets a positive return, and then both are insolvent only when they both get 0 returns. Suppose, for now, that bank 1 is fully invested in asset 1. Then bank 2 wants to choose the same portfolio if

$$\theta \frac{(1+s)\left[R_1 - d\right]}{1-s^2} > \theta^2 \frac{\left[R_2 + sR_1 - (1+s)d\right]}{1-s^2} + \theta(1-\theta)\frac{\left[R_2 - (1+s)d\right]}{1-s^2} + (1-\theta)\theta\left[\frac{s}{1-s^2}\left(R_1 - (1+s)d\right) - d\right].$$

[27] Note that $R_1 \geq R_2 > d$, so the right hand side is less than $R_2[1 - s^2(1-\theta) + s(1+s)(1-\theta)]$ which is $R_2[1 + s(1-\theta)]$, directly comparable to the left hand side.

This simplifies to

$$R_1 > R_2 - (1+s)(1-\theta)d.$$

Again, this is always true when $R_1 \geq R_2$, and holds even if $R_1 < R_2$, this time for large enough $(1+s)(1-\theta)d$.

Thus, there exist equilibria in a variety of settings in which both banks fully invest in a risky asset that is first order stochastically dominated by another because of the incentive to correlate their investment. There also always exist equilibria in which they both invest in the asset that pays the highest payoff, even if that fully correlates their portfolios.

More generally, the above analysis implies that if they can invest in different portfolios that have the marginal payoff distributions, and can choose whether to correlate them or not, then they will strictly prefer to correlate them. In such cases, correlation is the unique equilibrium.

If we look at the social value of investments, no costs of bankruptcy are born if the banks are both solvent, and so having the banks choose independent portfolios rather than highly correlated ones is generally preferable. Indeed, in the second case of the above example, the social optimum whenever R_2 and R_1 are close to each other is to have one bank invest in one asset and the other bank invest in the other (or to have banks hold both R_1 and R_2). Instead the banks prefer to hold all of the same investment. Note that this misalignment between bank incentives to correlate and what is socially optimal depends on how correlation affects the expected number of defaults in the network. In the second case, a single asset paying off is enough to ensure solvency of all banks, which in this setting is equivalent to assuming independent portfolios strictly decrease systemic risk. If on the contrary we look at the first case that assumes $s(R_i - d) < d$ for $i = 1, 2$—i.e., a bank defaults as soon as its own portfolio pays zero irrespective of the realization of its counterparty's—then correlated portfolios are no longer inefficient since they actually reduce systemic risk.

3.2.2 A General Result on Correlation and 'Risk Stacking'

To see how the above example generalizes, we now consider a set $N = \{1, \ldots, n\}$ of banks, whose financial interdependencies are summarized in the matrix of equity holdings $\mathbf{S}$ and debt holdings $\mathbf{D}$. Each bank has a proprietary investment opportunity that yields a return $R_i > 0$ with probability θ and 0 otherwise. We examine how they would choose to correlate their returns.

To examine the correlation in full generality, we model the world as having a large number of equally likely states of nature, and each bank can choose in which of those states they get $p_i = R_i$ and in which they get $p_i = 0$, subject to having a total probability of θ of getting $p_i = R_i$. We model this by introducing a set of K equally likely primitive payoff-irrelevant states, and a set of K Arrow-Debreu securities, each paying off in exactly one state. Each bank can then choose which states it gets 0 in and which ones it gets R_i in, as long as it maintains an expected return equal to θR_i; i.e., each bank chooses a fraction θ of the K

states that it wants to get R_i in.[28] An equilibrium is a state-contingent portfolio return for each bank that is feasible and optimal given equilibrium strategies of others in the financial network.[29]

Thus, if banks want perfectly correlated portfolios, they will all choose to get 0 in the same states, while to have independent portfolios they will all choose to get their respective 0's in a pattern that corresponds to a binomial distribution. They could also choose to get their 0's only when all others get R_{-i}, and thus negatively correlate their portfolios, and so forth.

Note that in this world, we can write the V_is as a function of the vector of 0s and Rs that are realized. Let $p_{-i} = \mathbf{R_{-i}}$ denote that all banks other than i have received R_js, and $p_{-i} = \mathbf{0}$ denote that all other banks have gotten 0s.

PROPOSITION **2.** *Suppose that all banks are solvent if they all get returns of R_i, that no bank in the network can remain solvent if it is the only one in the network with a nonzero portfolio realization (i.e., $V_i(p_i = R_i, p_{-i} = \mathbf{0}) \leq 0$ for all i), and that there is at least one bank i such that $V_i(p_i = R_i, p_{-i} = \mathbf{R}_{-i}) > V_i(p_i = 0, p_{-i} = \mathbf{R}_{-i})$. Then there is no equilibrium of the investment game in which portfolios are independent across banks, but there exists an equilibrium in which they are perfectly correlated.*

Note that the condition $V_i(p_i = R_i, p_{-i} = \mathbf{R}_{-i}) > V_i(p_i = 0, p_{-i} = \mathbf{R}_{-i})$ is extremely weak, and necessary to get any results. If it was true that all banks saw no increase in value when their portfolio changed, even in the best possible state in terms of the payments made by other banks in the network, then banks would not care about their returns at all - everything would be an equilibrium. That would require that essentially all returns $\sum_i R_i$ in the full network be owed as debt to outsiders, and so no bank in the network would ever have any equity value. The condition that $V_i(p_i = R_i, p_{-i} = \mathbf{0}) \leq 0$ for all i is also weak in that it applies to the extreme case where all other organizations in the whole economy get no return, and requires that the only bank with a positive return cannot survive. This condition is not necessary for the result (e.g., see the example above), but frees the proof from considering uninteresting subcases.[30]

[28]Here we cap how much that can invest in any state. Without that requirement, there are even more extreme equilibria in which banks fully correlate and invest even more in the risky asset - in fact they earn the highest return by putting all of their investments in just one state, which then minimizes the probability of having to pay any debt.

[29]We take θ to be a rational number, and K to be large enough so that $K\theta^n$ is an integer. Banks must choose θK different states to get R_i in, so they cannot, for instance, choose to get $2R_i$ in some states. As will become clear in the proof, the ideas generalize.

[30]Without that condition a bank may prefer not to correlate its portfolio from others if, by doing so, it can prevent the default of some of its debtors when $p_{-i} = \mathbf{0}$. Indeed, if these debtors happen to have a high enough equity claim on bank i and on its portfolio realization, they may remain solvent despite all other assets paying zero. This can be beneficial for i if it means getting a net debt coming in in such states. For independence to be an equilibrium, such incentive must hold for all banks. This is impossible since banks that are net borrowers cannot benefit from such feedback effect, and will always prefer defaulting as well when $p_{-i} = \mathbf{0}$. It also cannot be beneficial for privately owned banks: since no-one holds equity shares in them, they cannot prevent their debtors' default and hence cannot gain from uncorrelated portfolios.

Proposition 2 shows that fairly weak conditions are sufficient to ensure that full correlation is always an equilibrium and that independent portfolios are not part of any equilibrium. Intuitively the latter result comes from the fact that independence puts positive probability on states in which all banks but one get a zero portfolio realization, and that generally a bank will prefer to move its high portfolio realization from that state in which it would not receive it (being dragged down by the other banks) to a more favorable state in which it would remain solvent.

We refer to this incentive to correlate as '*risk stacking*' rather than risk shifting. Risk shifting is a somewhat different phenomenon in which an investor has an incentive to arrange a portfolio so that the risk falls on other investors, while here the phenomenon is to explicitly 'stack' all of the risk of all organizations into the same states. The result in Elliott, Georg and Hazell (2018) is an example of shifting, as banks want to correlate their assets to shift losses from states in which the bank is solvent—in which case the loss is incurred by *shareholders*—to states in which the bank defaults—in which case it is incurred by *debtholders*. The intuition behind our result is somewhat different, as our result holds even if we relax limited liability. The incentive to correlate here comes from the fact that high portfolio realizations are complements: a bank gains more by remaining solvent and getting $p_i = R_i$ when others also have high portfolio realizations since its own value depends positively on others' through financial interdependencies. In particular, as long as there exists an equity cycle in the financial network, then a bank's positive return gets magnified when others on the cycle remain solvent, which incentivizes correlation of portfolios even if banks do not act under limited liability. This intuition holds generally regardless of how bankruptcies are resolved or how large those costs are, and in particular without assuming that a bank would bear the costs of its counterparty's bankruptcy.

Risk Aversion Although we have worked with risk-neutral banks to keep the analysis uncluttered and to emphasize the impact of financial interdependencies, it should again be apparent that, just as in the risky investment case of Section 3.1, the above results extend to the case in which investors are risk averse. If others have correlated investments and when they become insolvent so will a given bank, then it is in that bank's interest to correlate its investments with those of the others regardless of its risk aversion. Thus, perfect correlation remains an equilibrium regardless of risk tolerances. In addition, the result that full independence is not an equilibrium holds for exactly the same reasons, regardless of risk tolerance.

3.2.3 The Inefficiency of Full Correlation

In terms of efficiency, maximizing the total value of all private investors in the economy is equivalent to minimizing the expected number of defaults. Indeed, the correlation structure of investments across banks does not change the expected aggregate portfolio value $\sum_i p_i = \theta \sum_i R_i$, but it does impact the set of defaulting banks and hence the amount of bankruptcy costs incurred. Correlated investments across banks are then socially efficient if and only if they induce a lower expected number of defaults than some other configuration. This holds if any bank that gets $p_i = 0$ always becomes insolvent irrespective of what else happens to other portfolios: independent investments do not attenuate systemic risk since high portfolio realizations from some banks can never prevent another from defaulting. In that extreme case, correlated investments are socially efficient, and banks' incentives are aligned with that of the social planner. However, as soon as correlation worsens contagion risk, the equilibrium is generally not socially optimal. This will be true in many cases of interest, such as when θ is high and so full independence would make it rare for banks to get low returns together and bankruptcies would be much rarer under independence than under full correlation. In general, full correlation is the worst possible case for bankruptcy costs since all banks are insolvent whenever their return is 0, and so every time someone gets a 0 outcome, they incur bankruptcy costs. If instead, one changes the correlation structure so that there are states in which some banks get 0 returns and do not become insolvent (and also so that the only insolvent banks are ones with 0 returns), then one decreases the bankruptcy costs. As long as the total frequency of overall insolvencies is less without perfect correlation, so that there are fewer bankruptcies than 0's on average (and presuming symmetry in bankruptcy costs), then bankruptcy costs drop when one moves away from perfect correlation. Thus, in most cases the decentralized equilibrium is socially inefficient, since Proposition 2 highlights that there does not exist an equilibrium in which banks choose independent portfolios, but there exists one in which they choose perfect correlation, irrespective of how correlation affects systemic risk.

3.2.4 Uniqueness of the Full Correlation Equilibrium

Proposition 2 provides insight into two of the most natural types of correlation, but does not address all possible correlation structures. Generally, banks have incentives to line up their positive returns, so as not to be dragged down by other banks. Nonetheless, one does need stronger conditions to ensure that all banks want to perfectly align their returns in *all* equilibria - such a strong form of uniqueness requires ruling out equilibria in which the network is partitioned into subsets of banks that correlate their portfolios within each subset, but choose (partially) uncorrelated portfolios across subsets. As we show next, under stronger conditions perfect correlation of portfolios is the unique equilibrium.

PROPOSITION **3.** *If the value from a high portfolio realization when others receive p_{-i} is increasing in the number of high portfolio realizations among other banks—i.e.*

$$V_i(p_i = R_i, p_{-i}) - V_i(p_i = 0, p_{-i}) > V_i(p_i = R_i, p'_{-i}) - V_i(p_i = 0, p'_{-i})$$

for each i, p_{-i}, p'_{-i} such that $|\{j \neq i : p_j = R_j\}| > |\{j \neq i : p'_j = R_j\}|$, then there is a unique equilibrium *out of all possible portfolio configurations and it involves perfect correlation.*

The sufficient condition that we give for perfect correlation to be the unique equilibrium is that the marginal gain in market value from a high realization of one's own portfolio is strictly increasing in the number of other banks that also have a high portfolio realization. This holds under a symmetry assumption on the underlying financial network, such that no bank would prefer to correlate with a particular counterparty as opposed to some other larger set of banks.

Without such a condition, there can exist other equilibria in which there is partial correlation of portfolios. Figure 2 gives an example of a financial network in which this condition does not hold, and describes an equilibrium in which banks correlate their portfolios within subgroups.

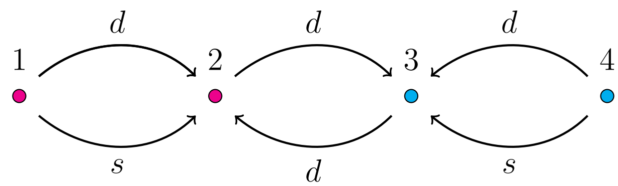


Figure 2: A network in which Proposition 3 does not apply. Bank 2 and 3 owe debt d to each other; bank 1 (4) owes d to 2 (3) and owns an equity share s of bank 2 (3) as well. Let $R < d < (1+s)R$ such that bank 1 (4) remains solvent if and only if both bank 1 (4) and 2 (3) have high portfolio realizations. Perfect correlation of portfolios is an equilibrium, but it is not the only one. There also exists an equilibrium in which banks in pink perfectly correlate their portfolio with each other, but do not correlate with the blue banks—that is, in the θK states in which $p_1 = p_2 = R$, $p_3 = p_4 = 0$ and reciprocally.

Note that even when there exist other equilibria with partial correlation, the banks each get their highest possible payoff in the full correlation equilibrium.

3.2.5 Oversight and Combating Incentives for Correlation

Given the result that at least some form of correlation in portfolios is to be expected in all equilibria of the financial system, then running stress tests for each bank separately overlooks a significant source of systemic risk. Indeed, without detailed information on the overall network, a bank-specific stress test does not capture the fact that a decrease in a bank's direct asset holdings is also likely to depress other banks' values, and hence also depress the value of its inter-bank assets. Going back to the example of the last financial crisis, a single bank stress-testing its own portfolio would underestimate the imminent collapse, as it would depress parts of the portfolio but not fully account for the fact that many other assets would be dropping in value at the same time due to the interlinkages, firesales, and multitude of network feedbacks. Better oversight and network-based stress testing could identify correlation patterns before they become catastrophic, given the many pressures for inter-linked financial organizations to invest in the same assets.

3.3 Too Few Partners: The Extensive Margin

In the previous sections, we looked at the intensive margin of investment choices, and highlighted the incentives to take excessively risky investments and to choose correlated investments. We now turn to the extensive margin and study how many counterparties a bank chooses. We find an under-diversification of bank portfolios in terms of number of partners.

The above results presume that banks have access to similar investments. There are various reasons, including regional presence, international boundaries, as well as proprietary advantages, that some banks might have access to investments that others do not. These, together with dynamic variations in banks' portfolios, can induce them to contract with each other–as evidenced by the large inter-financial interdependencies mentioned in the introduction.

As also mentioned in the introduction, one of the many things that make financial networks special is that financial contagion depends non-monotonically on the average degree (Elliott, Golub and Jackson (2014)). A higher average number of counterparties facilitates contagion conditional on a first failure as more organizations can be reached from the first failure. However, it also leads bank interdependencies to become more diversified through lower exposure to any single counterparty (holding fixed the total amount of overall exposure), reducing the risk of a first-failure and its probability of leading others into bankruptcy. Beyond a certain level of financial integration, this diversification effect dominates. There is thus a critical number of counterparties at which systemic risk is maximal, what Elliott, Golub and Jackson (2014) call the sweetspot. We now examine how banks choose the number of their counterparties.

3.3.1 Syndicated Investments

Let's consider a bank's choice between two different regimes: one in which it makes joint investments with $m+1$ other banks and another where it makes joint investments with m other banks. Each bank has debt liability $D_i^L = d$ to outside investors and no other contracts—all banks are hence symmetric. Without loss of generality, we write the problem from bank 1's perspective and consider partnering with the first m banks. It costs c to contract with each other bank.

A bank prefers to invest (equally) with m other banks if and only if

$$\mathbb{E}\left[\sum_{i=1}^{m+1}\frac{p_i}{m+1} - d \middle| \sum_{i=1}^{m+1}\frac{p_i}{m+1} \geq d\right] \Pr\left[\sum_{i=1}^{m+1}\frac{p_i}{m+1} \geq d\right]$$
$$-\mathbb{E}\left[\sum_{i=1}^{m}\frac{p_i}{m} - d \middle| \sum_{i=1}^{m}\frac{p_i}{m} \geq d\right] \Pr\left[\sum_{i=1}^{m}\frac{p_i}{m} \geq d\right] \geq c.$$

If the p_i's are perfectly correlated, then the left-hand-side expression is 0, and so there is no potential benefit to syndication. So, consider the case in which the p_i's are less than perfectly correlated.

The distribution $\sum_{i=1}^{m} \frac{p_i}{m}$ is a mean-preserving spread of $\sum_{i=1}^{m+1} \frac{p_i}{m+1}$, and so the expectation of a nondecreasing and convex function of these variables will be higher under m than $m+1$ (e.g., see Hadar and Russel (1971)). Since $\left[\sum_{i=1}^{m} \frac{p_i}{m} - d\right]^+$ is nondecreasing and convex, it follows that

$$\mathbb{E}\left[\sum_{i=1}^{m} \frac{p_i}{m} - d \middle| \sum_{i=1}^{m} \frac{p_i}{m} \geq d\right] \Pr\left[\sum_{i=1}^{m} \frac{p_i}{m} \geq d\right]$$

is decreasing in m, and so the banks prefer not to be involved in syndicates, and choose $m = 0$.

A planner who values the total value of all returns and costs in a society is not concerned with debt repayments which are simply transfers, nor the expected returns which are realized regardless of the solvencies. The planner is concerned with the contracting and bankruptcy costs. Noting that all of the banks would go bankrupt at the same time in this case, the planner's problem would then prefer swaps of size $m+1$ to m if and only if:

$$b\left(\Pr\left[\sum_{i=1}^{m+1} \frac{p_i}{m+1} \geq d\right] - \Pr\left[\sum_{i=1}^{m} \frac{p_i}{m} \geq d\right]\right) \geq c.$$

The sign of

$$\Pr\left[\sum_{i=1}^{m+1} \frac{p_i}{m+1} \geq d\right] - \Pr\left[\sum_{i=1}^{m} \frac{p_i}{m} \geq d\right]$$

depends on the size of d relative to the distribution of the p_i's. If sensible-enough investments from debt-holders are to be expected, returns should on expectation be enough to cover liabilities $\mathbb{E}[p_i] \geq d$. Then taking an average over m compared to $m+1$ observations is a mean preserving spread and lowers the probability that the average is above d. Thus, in most cases of interest we would expect $\Pr\left[\sum_{i=1}^{m} \frac{p_i}{m} \geq d\right]$ to be increasing and concave in m, and correspondingly the difference to be positive and decreasing in m.[31]

If the size of bankruptcy costs relative to partnering costs b/c is nontrivial, then the planner will prefer to have some partnering, while the banks will prefer to remain isolated—and so they underconnect relative to what is socially optimal.

The effect of asset correlation on the optimal number of partners is to reduce the benefits of partnering.

The above analysis focuses on the incentives for complete sharing. However, in that situation there are no possibilities of contagion. Next we turn to a setting with possibilities of defaults on payments, or drops in equity values of one bank, and its potential effect on others, which introduces additional externalities.

[31]This depends on the distribution. For instance, for the normal distribution and many other continuous unimodal distributions, this is true. But it can fail for some m for multimodal distributions. For instance, take a binomial distribution with equal likelihoods of 1 and 0's, and then set $d = 2/5$. the probability that the average of 1 draw is below d is $1/2$, for 2 draws it is $1/4$, but for 3 draws it is $1/2$ again, then for 4 draws it is $5/16$. It converges to 0, but has some nonmonotonicities in its convergence.

3.3.2 Equity Shares

Consider the problem of symmetric banks that can acquire a total equity share s in other banks, and must decide between how many banks to split this investment. Each bank has debt liability $D_i^L = d$ to outside investors, and no other financial contracts. For the sake of tractability we examine *cliques*, where a clique is a set of $m+1$ banks in which each bank owns a share s/m in every other member for some $m \geq 1$. For instance, banks could choose to form cliques of 3 banks, in which each acquires a share $s/2$ in the two others. Focusing on cliques is with loss of generality, but greatly simplifies the analysis and can be interpreted as a stylized representation of clustered networks.[32] Finally suppose that direct asset holdings are i.i.d. across banks, with $\mathbb{E}[p_i] > d$, so that there is a possibility of solvency. If a bank defaults, suppose that a bankruptcy cost of b is incurred.

Partnering with more banks decreases the variance in a bank's portfolio, which can improve its probability of solvency. When choosing the size of a clique, bank i solves

$$\max_m \quad \mathbb{E}[V_i(m)|V_i(m) \geq 0] \Pr[V_i(m) \geq 0] - cm.$$

Here

$$V_i(m) = \max[0, p_i - d + \frac{s}{m} \sum_{j \neq i} V_j(m)]$$

So, again, since averaging over a smaller number of V_j's will often be a mean preserving spread, and the max function is convex, this leads banks to prefer $m = 0$ and have lower expected values as m increases in this setting as well. Here, however, things are slightly more complicated given that the V_j's now have distributions that are correlated, and it is possible for the distribution not to be ordered by second order stochastic dominance as m changes - so there can be some nonmonotonicities. Thus, we provide some illustrative calculations for specific distributions below, and still find them to be decreasing in m and for banks to prefer $m = 0$.

Regarding the socially optimal level of m, given that the total value of all investments – the p_i's – are all realized regardless of any contracting and which banks are solvent, and all of the equity and debt are only transfers, the only part of the total value in a society that changes with m are bankruptcy and connection costs. Given the symmetry, the social optimum is the m that maximizes

$$-b \Pr[V_i(m) < 0] - cm.$$

This is equivalent to maximizing

$$b - b \Pr[V_i(m) < 0] - cm$$

[32] Clique-based networks can arise endogenously when banks are concerned with second-order counterparty risk—see Erol (2019) for a model of financial network formation in which equilibrium networks in the absence of regulation are composed of disjoint cliques.

or

$$b \Pr[V_i(m) \geq 0] - cm$$

We thus again find that the banks tend to prefer to have no partners, while for nontrivial values of b/c, the social optimum is some positive m.

We provide some calculations that illustrate the differences for various distributions and parameter values. Figure 3 depicts the objective of an individual bank and of the planner for different debt levels, assuming that bank portfolios are uniformly distributed. As argued above, it is always optimal for banks to choose not to hold any shares in each other. Indeed, although more counterparties do increase the probability that a bank remains solvent, that is not not enough to compensate the decrease in its expected value conditional on being solvent. The planner however prefers some diversification and incentives are misaligned.

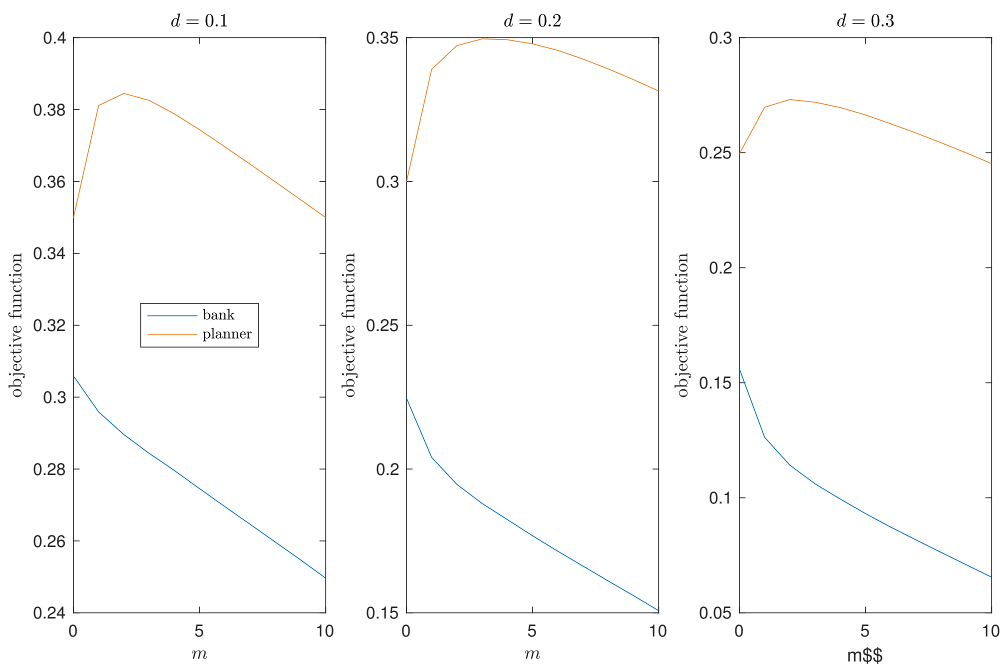


Figure 3: The bank and planner's objectives as functions of the number of partners of a given bank. The parameters for these calculations are $p_i \sim U(0, 0.8)$, $s = 0.5$, $b = 0.4$, $c = 0.005$, and three different levels of d are presented.

Figure 4 shows how the objective functions vary with the distribution of the p_i's, which affects the variance and tails of the distributions, and hence the probabilities of bankruptcies and those associated costs.

3.3.3 Discussion of Banks' Under-Investment in Partnerships

The intuition behind the inefficiently low number of partners chosen by banks can be seen as follows. In the above analyses, banks do not bear any of their own bankruptcy costs, nor

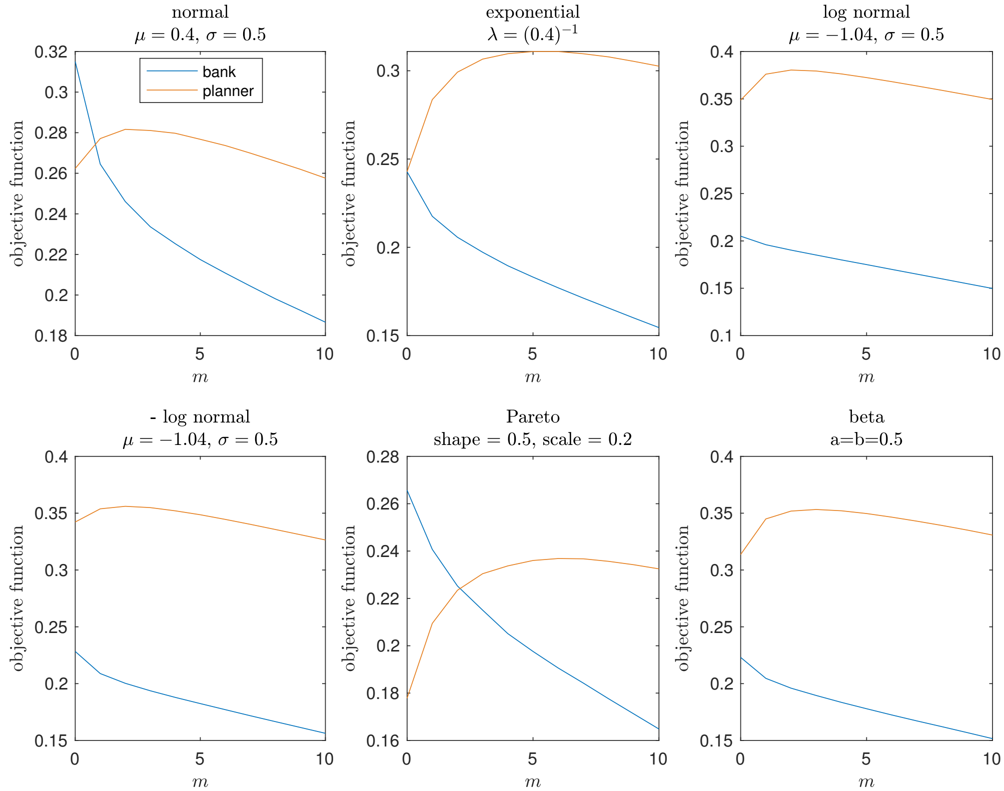


Figure 4: The bank and planner's objectives as functions of the number of partners of a given bank, for debt level $d = 0.2$, equity share $s = 0.5$, bankruptcy cost $b = 0.4$, and partnering cost $c = 0.005$. The results are presented for several different distributions of p_i, which are parametrized so as to keep the expected value of the asset constant to $\mathbb{E}[p_i] = 0.4$ (for the mirror-image log normal and beta distributions, this requires shifting the distribution appropriately).

do they bear each other's bankruptcy costs. In that context, consider adding some contract – either syndication or some other claims on each other – that would make a bank solvent when otherwise it would be insolvent. The injection of capital from other banks is costly to them, and all it does is then save an enterprise that ends up paying debts to outside investors that exceed the value of its own investments. So, the other banks essentially are paying something to add a net negative value to the overall value of all the banks. On average this *has* to be a net negative for the total value of all the banks. Thus, the results above are not special to a symmetric case: in any setting in which banks begin by only having debt liabilities to outsiders, additional partnering between banks that sometimes prevents the default of some must end up decreasing the total expected value of all banks involved. Even with asymmetries, some of the banks prefer not to have such partnering, and thus it would be blocked. Essentially, the banks do not bear the brunt of their bankruptcy costs, and so from an ex ante perspective, they prefer to fail when they are insolvent: spreading payoffs around to keep each other solvent only decreases their total expected value.

Sufficient risk aversion would change the incentives of the banks to partner with each other. With high enough levels of risk aversion, bank owners would be willing to shift some of their payoffs from situations in which they are wealthy to some of the situations in which they are insolvent. Then partnering with other banks would allow them to cross-insure, smoothing their expected payoffs across different portfolio realizations, and increasing their expected utility for high levels of risk aversion. Although sufficiently high levels of risk aversion would enhance bank owners' incentives for partnering, it would still not lead to full efficiency. These risk aversion effects would then also enter a social planner's calculations - who would evaluate the value of these cross payments in a similar manner since the planner values the expected utilities of the banks' owners. However, there would remain a key difference: the social optimum would still account for bankruptcy costs and lost debt payments to outside investors, while the banks' owners would not. Thus, they would still have lower incentives to avoid insolvencies than what would be socially optimal.

If banks have some cross debt holdings – so for some other reason they end up having debt in each other (e.g., for meeting short-term reserve requirements, etc.), then they do bear some of each other's bankruptcy costs and have more incentives to cross insure. Nonetheless, as long as some of their debts are to outsiders, they are not facing the full costs of their insolvencies. So, as long as they take deposits and have debts to outside investors, their calculations of the values of their portfolios omit key bankruptcy costs that impact other investors, and so they have incentives to take on too much risk, whether it be via excessively risky portfolios, under-insuring via partnering with each other, and over-correlating their investments whenever they do have partnerships.

4 Evaluating and Addressing Systemic Risk: Flying Jets without Instruments

As should be obvious by now, properly assessing systemic risk involves a holistic view of the network.[33] Contagion can occur directly between counterparties, or even indirectly without intermediate defaults as banks' values are interdependent even at a distance via chains of relationships.

An important component of systemic risk assessment is stress testing, which is usually run in a decentralized manner. The main input into many stress tests is balance sheet data, which describes the amount of each type of financial assets and liabilities held by each bank.

[33] Attempting to assess systemic risk without detailed and comprehensive network information is what Jackson (2019) refers to as "flying jets without instruments." It is assessing a highly complex interactive system without the necessary measurements and information to control it. Even though some stress tests and measures (e.g., S-risk) that can be built without network information may correlate with more precise full network measures, if those measures are only approximately capturing the real risks, it would be like flying a jet while only knowing altitude plus or minus a thousand feet or more, or airspeed plus or minus twenty percent, both of which could be catastrophic.

Depending on the jurisdiction, balance sheet data does not always provide complete, or even partial, information about the identity of one's counterparties, nor their other investments and counterparties, and hence about the network structure. Even if a stress test accounts for direct counterparties, it may miss information about systemic issues such as potential cascades or other indirect effects. Here, we argue that any measure of systemic risk based on "local" data can completely miss which banks are most likely to start a default cascade, or be caught up in one. The following example illustrates this point. The point is straightforward, but worth emphasizing given its importance.

We first illustrate this point with an example, and then discuss issues of ensuring solvency.

4.1 The Necessity of Network Information

For simplicity, consider a network in which banks only have debt contracts between each other. A measure of systemic risk based on local information only depends on $(D_i^A, D_i^L)_{i \in N}$. To show why this is insufficient information, we give an example of financial network in which two banks have *identical* balance sheets, and yet their defaults have significantly different consequences. Hence if the central authority were able to bailout one (and only one) of the two organizations, it could not take the optimal decision based on such local information. Consider the network composed of four banks depicted in Figure 5.

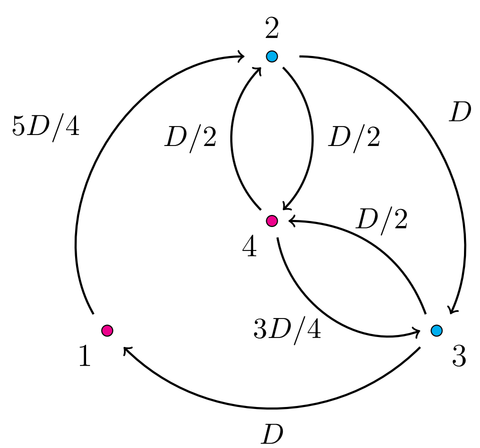


Figure 5: Arrows point in the direction that a debt is owed. Banks 1 and 4 (magenta) have total debt liabilities of $5D/4$ and debt assets of D, and are net debtors. Banks 2 and 3 (blue) have debt liabilities of $3D/2$ and debt assets of $7D/4$, and are net creditors.

Suppose the portfolios of Bank 1 and 4 yield 0, so that they are both insolvent, whereas Banks 2 and 3 earn returns on their investment between $3D/4$ and D. The recovery rate on assets of a defaulting bank is zero. Note that Bank 1 and 4 have the same balance sheet since $D_1^A = D_2^A = D$ and $D_1^L = D_2^L = 5D/4$. However, only Bank 1 induces widespread default contagion if it remains insolvent. Indeed, Banks 2 and 3 have enough buffer to absorb the shock of Bank 4's default, but not that of Bank 1. Hence, bailing out Bank 1 prevents the whole system from insolvency, while bailing out Bank 4 does not change anything and a full systemic failure occurs.

This example also highlights the fact that, without network information, one cannot even identify which banks are at risk of insolvency. For instance, if one examines the books of Bank 3 without knowing that Bank 2 is exposed to Bank 1, even if one knows the portfolio realizations of 3's counterparties, but does not know the looming failure of Bank 1 (which is not one of Bank 3's direct counterparties), it would appear that Bank 3 was free from danger of insolvency.

The first assessments of systemic risk that involve a nontrivial portion of the actual network are beginning to emerge, at least in Europe. For example, the European Central Bank has information on the counterparties involved in the largest exposures of most banks within its jurisdiction. This permits the construction of a network of a portion of the assets and liabilities within the European banking sector, and some pointers to banks outside of Europe, and thus some of the first calculations of systemic risk of a nontrivial part of the network are beginning to emerge (e.g., see Covi, Gorpe and Kok (2018)). Similarly, the Bank of England has regulatory data on bilateral transactions between UK banks, allowing for the analysis of the UK interbank network in different asset classes (see Ferrara et al. (2017); Bardoscia et al. (2018)). This is an important move of the assessment of systemic risk in the right direction, but much more is needed and especially outside of Europe and for the growing shadow banking system which falls outside of most jurisdictions.

4.2 Financial Centrality

We now provide a network-based measure of financial impact of a given organization. Conceptually, given our approach, there is a unique and clear way to assess financial impact. What limits its implementation is a lack of regulation requiring all counterparties to be revealed to a central bank or other oversight agency. The actual computation of the measures below would be demanding, but not infeasible to accurately approximate, especially once such a system would be in place and could be constantly updated.

Within our model, the obvious way to define a bank's impact on the rest of the economy following a change in its portfolio is to calculate its net impact on the overall value in the economy, as follows. Let the *financial centrality* of i at some vector of portfolios $\mathbf{p}$ and network $(\mathbf{D}, \mathbf{S})$ from a change to p_i' for i be

$$FC_i(\mathbf{p}, p_i'; \mathbf{D}, \mathbf{S}) = \sum_{j:private} \left(V_j(\mathbf{p}) - V_j(\mathbf{p}_{-i}, p_i')\right).$$

This is the total impact on the economy (i.e., on the ultimate shareholders and debt-holders) that comes from a change in i's portfolio.[34] Note that this is equivalent to

$$FC_i(\mathbf{p}, p_i'; D, S) = p_i - p_i' - \sum_j \left(b_j(\mathbf{V}(\mathbf{p}), \mathbf{p}) - b_j(\mathbf{V}(\mathbf{p}_{-i}, p_i'), \mathbf{p}_{-i}, p_i')\right),$$

which counts the total change in the portfolio and the total incidence of all changes in bankruptcy costs.

We define the *net financial centrality* of i as[35]

$$NFC_i(\mathbf{p}, p_i'; D, S) = \sum_{j:private} (V_j(\mathbf{p}) - V_j(\mathbf{p}_{-i}, p_i')) - (p_i - p_i' - (b_i(\mathbf{V}(\mathbf{p}), \mathbf{p}) - b_i(\mathbf{V}(\mathbf{p}_{-i}, p_i'), \mathbf{p}_{-i}, p_i'))\,.$$

This is the impact beyond the direct drop in i's portfolio value, and its own induced bankruptcy. Hence this solely captures the cascade that i causes in the economy from a change in the value of its portfolio. If there are no bankruptcies, or if only i goes bankrupt, then the net financial centrality of i is 0. As an easy extension, just as stress testing is done with certain percentage drops in asset values, these measures can be done in the same way, and one can find a distribution of these measures as one varies p_i'.

With appropriate network information, one can also do a systemic stress test due to a change from portfolio values p to p' by examining

$$\sum_{j:private} (V_j(\mathbf{p}) - V_j(\mathbf{p}'))\,.$$

In a world with only debt between banks, these measures capture chains of cascading defaults. Some banks can stop such cascades if they have enough value and/or small enough debt liabilities compared to debt assets, $D_i^A \geq D_i^L$. Note also that these "chains" could hit some banks multiple times and so intersect. With equity, cascades do not follow direct default chains but can skip a bank – that is, bank k could own j who owns i. Even if j does not default, its value could go down if i defaults, which could indirectly cause k to default. Hence with equity we can have indirect failures, whereas with debt there must only be chains of direct failures.

Regardless of the presence of equity, debts are still a key ingredient since they drive defaults. For example, in a balanced network, in which debt assets exactly compensate debt liabilities on every bank's balance sheet, no bank has any net financial centrality when considering the best equilibrium. Indeed, a bank is always able to repay its debt assuming all its counterparties are solvent and portfolios are non-negative. However such a network

[34] Changes in the values of public companies eventually all indirectly accrue to private equity and debt holders, and so including any of the public values would amount to double counting.

[35] Implicit in defining financial centrality, one has to take a stand on which equilibrium set of values is being used since those define the values $V(\mathbf{p})$ and thus the bankruptcy costs $b(\mathbf{V}(\mathbf{p}), \mathbf{p})$. Typically we are interested in either the best or worst equilibrium, but one could make other choices, or change from best to worst if one anticipates a freezing of payments in response to the failure of some organization(s). For some discussion about the importance of uncertainty about which equilibrium applies, see Roukny et al. (2018).

could still be very fragile. For instance, consider a balanced network with no contracts other than debt. One can change banks on paths away from i to have D_j^L slightly higher than D_j^A. If debts are large relative to the p's, then the resulting default centralities can be very large. More generally, this implies that (net) financial centralities are discontinuous and can be very sensitive, especially to the debt structure in the network.

4.3 The Size of Necessary Bailouts

The possibility of a repayment cascade in a network can be exploited by a regulator who tries to minimize default costs. Just as defaults cascade, the same operates in reverse and a well-placed bailout can have far-reaching consequences. This problem relates to Demange (2016), who characterizes the optimal cash injection policy in a network of financial liabilities under proportional rationing in case of default. She defines a threat index that identifies banks with highest marginal social value of liquidity, assuming the policy does not change the set of defaulting banks. Here, instead, we examine how much of an injection is needed to avoid any cascades and defaults, as a function of the network and portfolios.

Consider a regulator that can inject liquidity into the network, to ensure that some banks remain solvent—i.e., it can bailout a subset of banks $B \subseteq N$ by changing their portfolio values from p_i to some $p_i' > p_i$ by making direct payments to the banks of $p_i' - p_i$. If bank $i \in B$ is sufficiently bailed-out, it pays back its debt to all its creditors, who then may become solvent themselves. We now investigate the necessary costs and how the banks that need to be bailed out depend on the cycles in the network. We also highlight the difference between the best and worst equilibria.

4.3.1 Balanced Networks

It is useful to begin by examining some benchmark cases that we refer to as *balanced networks*, as they play key roles in the more general characterization.

For the sake of Proposition 4, we presume that there are no cross-equity holdings, and return to describe that case later. Furthermore, all the definitions that follow are relative to some specification of $\mathbf{p}, \mathbf{D}$, and we omit its mention.

We say that a network is *weakly portfolio balanced* if $p_i + D_i^A \geq D_i^L$ for all i. This is a weaker requirement than debt-balance, just requiring that each bank's (non-equity) assets are enough to cover its debt liabilities, presuming its incoming debt assets all are fully valued.

We say that a network is *exactly portfolio balanced* if $p_i + D_i^A = D_i^L$ for all i. Since all debt claims and liabilities cancel out on aggregate, exact balance implies $\sum_i p_i = 0$, and given nonnegative portfolios then implies that $p_i = 0$ for all i. This then also implies a network has exactly balanced-portfolios only if it has balanced debt: $D_i^A = D_i^L$ for all i. Under this requirement, a bank is able to meet its debt obligations going out if and only if it receives all the debt payments it has coming in.

Any of these forms of balance is sufficient for all organizations to be solvent in the best equilibrium. This follows since if all banks but i honor their debt contracts then i can also pay back its debt fully in a weakly balanced network. Essentially, all debts can be canceled out, if one allows strings of canceling. This logic holds irrespective of the network structure.

Things are quite different in the worst equilibrium, as it may be that some banks cannot pay out their debts unless some of their debt assets are paid in full. Hence it can be that if some other banks are insolvent, then bank i will be as well - and these insolvencies can cascade and cycle, as in the following example.

A Wheel Example. Consider a balanced-wheel network composed of n banks, such that bank i owes debt D to bank $i+1$. To close the wheel, bank n owes debt D to bank 1. Note that even when all portfolios yield zero, there exists a best equilibrium in which all banks remain solvent - the debts are all canceled out against each other. In contrast, when the banks' investments are worth less than D, then the worst equilibrium has all banks insolvent. The worst equilibrium corresponds to a situation in which a bank cannot pay out its debt until it receives sufficient payments in.

In this network, bailing-out a single bank by giving it D is enough to have all payments made and having all banks be solvent in the worst equilibrium. For instance, if someone bails out bank 1 by injecting D units of liquidity in this bank, then 1 pays back its debt to 2, which is then able to repay bank 3, and this cascades up to bank n who repays bank 1, who can then repay whomever bailed it out. The regulator recoups its initial liquidity input at the end of the cascade, and is thus able to prevent n defaults at essentially no cost.

In cases in which the worst equilibrium differs from the best equilibrium, one can interpret the worst equilibrium as a coordination failure since all banks could have written-off their counterparties' debt without cost so as to avoid a general default of the system. In practice banks are unlikely to be able to coordinate in such a way if debts involve cycles rather than just direct canceling between two counterparties, as it would require all write-offs to be done simultaneously to maintain solvency.[36] Moreover, in practice the debts have different maturities and other covenants and priorities that further complicate any canceling out without an economy-wide renegotiation. This makes looking at the worst equilibrium important, and so we discuss both equilibria in what follows.

The worst equilibrium can be thought of as the requirement that a bank can pay back its debts if and only if it already has sufficient capital to cover all of its debts based on the amount of incoming debts that have already been paid to it, together with any bailout payments. In particular, this rules out partial payments: even if a bank has some money coming in, it cannot use that money to pay some of its debt until it is fully solvent. Such requirement makes sense if all debt claims have equal priority, but makes bailing-out more demanding: the minimal set of banks that need to be bailed-out can be strictly larger under this rule than when partial repayments are allowed. To see which banks are solvent amounts

[36]Some systems for such canceling are emerging, such as enterprises that offer "compression" services, which are essentially canceling out of cycles of contracts (e.g., see D'Errico and Roukny (2019)).

to running the algorithm that finds the worst equilibrium assuming bankruptcy costs that preclude any repayments. That is, first banks that are solvent without any incoming debt payments repay their debts, then given those debt payments there may be new solvent banks just based on their portfolios and these first debt payments, which then repay their debts, potentially leading to new solvencies, and one iterates and ultimately finds a set of solvent and insolvent banks.

In the above wheel example, bailing-out a single bank is enough to have a repayment cascade through the entire network. If banks are in a more complex network, the minimal set of banks that need to be bailed out in order to ensure that the system is fully solvent in the worst equilibrium can be more complicated. The optimal strategy for the regulator is to exploit cycles in the network, so as to induce as many cascades of repayments as possible. We now provide a characterization.

Let us say that a bank i is *unilaterally solvent* if $p_i \geq D_i^L$. This means that regardless of whether any of the other banks pay the debts that they owe to i, i is still able to cover its debts.

We say that a set of banks S is *iteratively strongly solvent* if it consists of a nonempty subset of banks $S_1 \subset S$ that are unilaterally solvent; and then iteratively sets S_k such that the banks $i \in S_k$ are solvent if they receive the debts from all banks in sets $S_1, \dots S_{k-1}$, but not if they only receive the debts from banks in $S_1, \dots S_{k-2}$:

$$p_i + \sum_{j \in S_1 \cup \cdots S_{k-1}} D_{ij} \geq D_i^L > p_i + \sum_{j \in S_1 \cup \cdots S_{k-2}} D_{ij}.$$

Note that if N is iteratively strongly solvent, then all organizations are solvent in the worst equilibrium. Proposition 4 provides weaker conditions that are necessary and sufficient for this to hold. This then provides a base to understand minimal bailouts.

Proposition 4. *All organizations are solvent in the best equilibrium if and only if it is weakly portfolio balanced.*

All organizations are solvent in the worst equilibrium if and only if it is weakly portfolio balanced and there exists an iteratively strongly solvent set that intersects each directed (simple) cycle.[37]

An implication of Proposition 4 is that in a weakly balanced network, if one has an iteratively strongly solvent set that intersects each directed (simple) cycle, then that implies that the whole set of banks is iteratively strongly solvent. This is the crux of the proof.

The proposition is less obvious than it appears since an insolvent bank can lie on several cycles at once, and could need all of its incoming debts to be paid before it can pay any out. Solvent banks on different cycles could lie at different distances from an insolvent bank, and

[37] A simple cycle is one that only contains any organization at most once. If there is a solvent bank on each simple cycle then there is one on every cycle, since every cycle contains a simple cycle. A simple cycle is a list of links $i_0 i_1$, $i_1 i_2$, $\dots i_K i_0$ such that $D_{i_k i_{k+1}} > 0$ for all $k = 0, \dots, K$ (with $K + 1 = 0$), and i_0 is the only organization that appears twice.

showing that each bank eventually gets all of its incoming debts paid before paying any of its outgoing debts is subtle. The proof is based on how directed simple cycles must work in a weakly balanced network and appears in the appendix.

Proposition 4 also implies that if these conditions are satisfied, then there is a unique equilibrium. Conversely, if portfolios are weakly balanced and there is no iteratively strongly solvent set intersecting every cycle, then there are necessarily multiple equilibria. Thus, in a weakly portfolio-balanced network (excluding equity), there is a unique equilibrium *if and only if* there exists an iteratively strongly solvent set that intersects each directed (simple) cycle.

One way to ensure having an iteratively strongly solvent set intersecting each directed cycle is to have at least one unilaterally solvent bank on each cycle, but this is not generally necessary, and so the iterative solvency condition is important. A special case is when the network is exactly portfolio-balanced, such that no bank has a capital buffer: then it becomes necessary to have at least one unilaterally solvent bank on each cycle for the whole system to clear in the worst equilibrium. That is because in the case of exact balance, each bank is solvent if and only if it receives all the debt payments it is owed. However, note also that with exact balance, no bank that has any debts can be unilaterally solvent since their debts in are just enough to cover their debts out and so they have no buffer to make them solvent without those debt payments coming in, and so in that case bailouts will be necessary to get to solvency.

Proposition 4 highlights both the difference between the best and worst equilibria, as well as the role of cycles in clearing interbank liabilities. It resonates with the increasingly used technique of portfolio compression, which allows banks to eliminate offsetting obligations with other organizations taking part in the process, exploiting cycles in the financial network (e.g., see D'Errico and Roukny (2019)). Portfolio compression then reduces gross interbank exposures while keeping net exposures constant, which not only reduces systemic risk but regulatory requirements of participants as well. Portfolio compression improves the worse equilibrium since less it lowers the debts that banks owe and thus reduces the amounts that they need to be solvent either unilaterally, or once receiving some incoming payments.

4.3.2 Minimum Bailouts and Imbalanced Portfolios

We now investigate the minimum bailouts needed to ensure solvency of all banks in both the best and the worst equilibrium.

The best equilibrium is relatively easy to understand. If the network is not weakly portfolio balanced then some banks must be defaulting, and each bank that is not weakly balanced needs bailouts to be brought back to solvency. It follows then from Proposition 4 that the minimum amount of capital that needs to be injected in the system to ensure its solvency is exactly

$$\sum_i [D_i^L - D_i^A - p_i]^+.$$

The worst equilibrium is more complex, as weak balance is necessary, but not sufficient for solvency. One needs the above minimum payment, but then one also needs to inject enough additional capital into some set of banks to ensure the existence of an iteratively strongly solvent set that intersects each directed cycle in the network.

This is summarized in the follow corollary to Proposition 4.

Corollary **1.** *Consider a possibly imbalanced network with no equity.*

(i) The minimum necessary bailout needed to ensure solvency of the entire network in the best equilibrium is the total net imbalance in the economy $\sum_i [D_i^L - D_i^A - p_i]^+$ *(which is 0 if the network is weakly portfolio balanced).*

(ii) The minimum necessary bailout needed to ensure solvency of the entire network in the worst equilibrium is the total net imbalance in the economy and *injecting the minimum additional capital to generate an iteratively strongly solvent set that intersects each directed cycle in the network.*

(iii) If the network is fully compressed (so that all cycles of debt in the network have been cleared), then the best and worst equilibria coincide and the minimum necessary bailout needed to ensure solvency of the entire network in the best equilibrium is just the total net imbalance in the economy.

To ensure solvency of the entire network, the regulator first has to ensure weak balance of the portfolios of all banks. That is necessary and sufficient to ensure solvency in the best equilibrium. Then we are also back to the logic of the weakly balanced case, and to ensure solvency in the worst equilibrium, the additional capital that is needed is the minimum that then generates an iteratively strongly solvent set intersecting each directed cycle.

We remark that the payments made to regain weak balance, $\sum_i [D_i^L - D_i^A - p_i]^+$, will not be recovered by the government or other entity that intervenes in the bailout. However, all the additional payments made in the case of the worst equilibrium can be recovered. For instance, as we saw in the wheel example, the payments that are made in a cycle all cancel out and can be recovered once the capital has cycled through the entire network. This logic holds more generally, even in more complex networks. Indeed, once necessary payments to regain weak balance has been made, each bank's balance sheet satisfies $p_i + D_i^A \geq D_i^L$. To ensure solvency in the worst equilibrium, an appropriate set of banks then needs to bailed-out, that is an additional $D_i^L - p_i$ needs to be injected in a subset of banks. However, since this guarantees solvency of the whole network, we know that in the end these banks will get their debt payments $D_i^A \geq D_i^L - p_i$ in full. Hence the regulator will be able to recover the additional capital it had to inject to generate the iteratively strongly solvent set.

Figure 6 illustrates some of these ideas. In the imbalanced case on the right, the net payment that is needed to reach weak balance is to give d to Bank 3, which is never recovered, and that is enough to ensure full solvency in the best equilibrium. Then paying an additional d to Bank 3 is the minimum additional bailout needed to ensure solvency in the worst equilibrium, as that is enough to get Bank 3 to make its payments, which then makes Bank 2 solvent, which then makes Bank 1 solvent.[38] Bank 2 being solvent means it is then able to repay its debt of d to Bank 3, which can be recovered by the entity that intervened. Hence net bailout costs here equal d, which is the payment needed to make the network weakly balanced.

Figure 6: Suppose all banks have $p_i = 0$. Ensuring solvency of the full network in the worst equilibrium in the balanced case (left panel), requires bailing out at least one bank per cycle, which can be done by either bailing out banks 1 and 3, or just bank 2. In the imbalanced case (right panel) the whole system clears if bank 3 is bailed out, and hence bailing out one bank per cycle is not necessary, and $\{3, 2, 1\}$ (in that order) form an iteratively strongly solvent set (once 3 gets an bailout injection of $2d$).

This problem of generating an iteratively strongly solvent set that intersects each directed cycle in the network is well-defined, but finding a minimum-cost set can be computationally hard. This is true both in terms of requiring a lot of information on the network structure, as well as the computational complexity of finding the minimal combination of banks to bailout to have at least one on each cycle. For instance, many banks will sit on multiple cycles, and even simply identifying all of the cycles in a network is challenging. This problem is in fact NP-hard.[39]

In practice, the problem is simplified for two reasons. First, many organizations will already be solvent, and it may only be a subset that are problematic—attention can then be concentrated on a subnetwork and some key organizations. Second, a core-periphery network (which many financial networks resemble) has a large amount of structure to it that makes its cycles easier to identify: it consists of a core clique together with a bunch of extra links to banks that have few connections to the core.

It can be significantly costlier to ensure solvency of an imbalanced than a balanced financial network, since the regulator necessarily has to inject the amount of net imbalance of all net borrowers. In reality, this imbalance can be large as many organizations have some debt contracts with partners who are private individuals and who are not otherwise involved in the network: for instance they have loans out as mortgages, or deposits that can be treated like debt for our purposes (e.g., demand deposits, certificates of deposit, overnight

[38] Paying $d/2$ to 1 would not be enough to get 2 to be solvent since it still gets no payments from 3 in the worst equilibrium. One would need to pay $3d/2$ to 2 in order to ensure that it would be unilaterally solvent. So, 3 is the cheapest option.

[39] Indeed, when looking for all simple cycles in a graph, one also solves the Hamiltonian cycle problem, which is known to be NP-complete. For an early algorithm to find all cycles, see Johnson (1975).

loans, money market accounts, etc.). These debts do not recycle into the network and so cannot be canceled out. Even though debts fully balance in aggregate—for each lender there is a borrower—some organizations or individuals are net lenders and others are net borrowers. Any organization belonging to the later category can be a first failure, and start a default cascade. They can also propagate failures, and are hence the critical organizations. On the contrary net lenders can never be the first to fail, but can be brought to bankruptcy as well if enough of their counterparties default.

Proposition 4 and Corollary 1 do not address the case in which banks hold equity in each other. That further complicates the calculations, since instead of just $[D_i^L - D_i^A - p_i]^+$, the imbalance of a bank now also depends on the value of its equity holdings in other banks, which then depends on who is solvent.

The general program to ensure solvency at minimum cost can be written as

$$\min_{p' \geq p : \mathbf{V}(\mathbf{p}') \geq 0} ||\mathbf{p}' - \mathbf{p}||,$$

where the $\mathbf{V}$ is chosen to be either the best or worst equilibrium, depending on which is of interest. Again, this requires that the imbalance is at most 0 for all organizations, but now this includes equity values and so has to be solved as a fixed point.

The algorithm for finding the amount needed to remove the net imbalance in the case of the best equilibrium is straightforward to describe. It is as follows.

Let

$$p_i^n = p_i + D_i^A - D_i^L.$$

Now, calculate the best equilibrium values associated with asset returns $\mathbf{p}^n$, equity holdings $\mathbf{S}$ (noting that equity values in negative-valued enterprises are 0), and no debt $\mathbf{D} = \mathbf{0}$. The opposite of the total sum of the valuations of the organizations with negative values (ignoring bankruptcy costs) is the minimum bailout that is needed. Effectively, we know that all debts will be repaid in a bailout that ensures full solvency, and then the resulting bank values will be the basis on which equity values accrue. Organizations that are still negative, including all of their equity positions are the ones that will require bailout payments.

In the case of the worst equilibrium, the same logic applies, but then the base values are associated with the worst equilibrium. Then once those payments are made, one *recalculates* the worst equilibrium values given those payments, but with the original $\mathbf{D}$. By doing this, one identifies banks that are then unilaterally solvent (after the initial bailout payments), as well as any resulting iteratively strongly solvent set by consequence of those unilateral solvencies. If these are not enough to intersect each directed cycle, then additional bailouts will be needed, and an algorithm needs to be run to find the cheapest set. Note that those bailouts might not even be used to generate unilateral solvencies, but might just be enough to generate secondary solvencies given the unilateral solvencies, which eventually generate more solvencies. This is the analog of the problem without equity, but just augmented by additional value calculations that include equity of the resulting solvent organizations for each possible configuration of bailouts that is considered and the corresponding worst equilibria. Here, part (iii) of the Corollary becomes particularly important, since it means that if one can compress the network, then the issues with the worst equilibrium are avoided and one only has to deal with the initial bailouts needed to restore weak balance, which are necessary in any case.

5 Concluding Remarks

We have highlighted two main points.

One is that to properly assess systemic risk one needs detailed network data. This one is "easy" to fix, as once one has counterparty information, although data-intensive, the way in which one should assess systemic risk is straightforward. We put "easy" in quotes since, although what is needed is simple and obvious, it may be politically difficult to get. Financial organizations, for a variety of reasons, would prefer to keep their detailed investment information private. It also opens questions of how public one makes the outcomes of such stress tests and how one acts upon the information. Nonetheless, it is clear that operating without such information is just asking for another financial crisis to happen, or else requires having excessively onerous regulation to ensure solvency regardless of the network conditions.

The second point is that the externalities in financial networks lead to several incentive problems:[40] organizations have incentives to take overly risky positions, to involve too few counterparties, and to overly correlate their portfolios with those of their counterparties. These are harder to fix. Excessive risk can be partly, but imperfectly, addressed by reserve requirements and/or bailouts as we have shown. The imperfection relates to the fact that such reserves are generally only imposed based on a portion of the liabilities and only for a subset of financial organizations (e.g., missing much of the shadow banking system). Incentives to take on too few counterparties and to overly correlate portfolios are also issues that have been ignored by policymakers, and not ones for which there are easy policies. Requiring that some markets have Central Counterparty Clearing Houses – CCPs – can be thought of as part of a solution to these issues. These pass all transactions through a central intermediary, or a few, which can monitor positions and impose margin requirements. One then has to worry about providing the CCPs with appropriate incentives and worry about their size.[41] Large government-sponsored enterprises that process huge amounts of securities have an uneven history of success, especially if one examines Fannie Mae and Freddie Mac's failures in the 2008 crisis. Moreover, although it can mitigate some of the systemic ramifications of the inefficiencies, it still does not eliminate the excessively correlated and

[40] There are other endogeneity issues that we have not discussed, for instance, whether two organizations wish to merge, or how large they become. In Appendix C we briefly discuss how banks' size interacts with systemic risk and investment incentives; but we leave further analysis to future research.

[41] See for instance, Duffie and Zhu (2011).

risky portfolios that are induced, and hence the individual bankruptcy costs that are still not incorporated.

Regardless of the precise policy that one undertakes, developing and maintaining a more complete picture of the network, and the portfolios of banks together with those of their counterparties, is a necessary first step both to improving crisis management and to better understanding and monitoring incentive distortions.

References

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "Systemic Risk and Stability in Financial Networks," *The American Economic Review*, 2015, *105* (2), 564–608.

Acharya, Viral V and Tanju Yorulmazer, "Too many to fail?An analysis of time-inconsistency in bank closure policies," *Journal of financial intermediation*, 2007, *16* (1), 1–31.

_ , Sreedhar T Bharath, and Anand Srinivasan, "Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries," *Journal of Financial Economics*, 2007, *85* (3), 787–821.

Admati, Anat R. and Martin F. Hellwig, *The Bankers New Clothes: Whats Wrong with Banking and What to Do about It*, Princeton University Press: Princeton, 2013.

Allen, Franklin, Ana Babus, and Elena Carletti, "Asset Commonality, Debt Maturity and Systemic Risk," *Journal of Financial Economics*, 2012, *104* (3), 519–534.

_ **and** _ , "Networks in finance," in Paul R. Kleindorfer and Yoram Gerry, eds., *The Network Challenge: Strategy, Profit, and Risk in an Interlinked World*, Wharton School Publishing, 2009, pp. 367–382.

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

Alvarez, Fernando and Gadi Barlevy, "Mandatory Disclosure and Financial Contagion," *Paper No. w21328, National Bureau of Economic Research*, 2015.

Arya, Anil and Jonathan Glover, "Excessive intervention, the collusion problem, and information system design," 2001.

Babus, Ana, "The formation of financial networks," *The RAND Journal of Economics*, 2016, *47* (2), 239–272.

— **and Tai-Wei Hu**, "Endogenous intermediation in over-the-counter markets," *Journal of Financial Economics*, 2017, *125* (1), 200–215.

Bardoscia, Marco, Gerardo Ferrara, Nicholas Vause, and Michael Yoganayagam, "Full Payment Algorithm," *Available at SSRN*, 2019.

— **, Ginestra Bianconi, and Gerardo Ferrara**, "Multiplex network analysis of the UK OTC derivatives market," *Staff Working Paper: Bank of England*, 2018.

— **, Stefano Battiston, Fabio Caccioli, and Guido Caldarelli**, "Pathways towards instability in financial networks," *Nature Communications*, 2017, *8*, 14416.

Barucca, Paolo, Marco Bardoscia, Fabio Caccioli, Marco D'Errico, Gabriele Visentin, Stefano Battiston, and Guido Caldarelli, "Network valuation in financial systems," *Available at SSRN 2795583*, 2016.

Bech, Morten L and Enghin Atalay, "The topology of the federal funds market," *Physica A: Statistical Mechanics and its Applications*, 2010, *389* (22), 5223–5246.

Billio, Monica, Mila Getmansky, Andrew W. Lo, and Loriana Pelizzon, "Econometric measures of connectedness and systemic risk in the finance and insurance sectors," *Journal of Financial Economics*, 2012, *104(3)*, 535–559.

Blasques, Francisco, Falk Bräuning, and Iman Van Lelyveld, "A dynamic network model of the unsecured interbank lending market," *Journal of Economic Dynamics and Control*, 2018, *90*, 310–342.

Branch, Ben, "The costs of bankruptcy: A review," *International Review of Financial Analysis*, 2002, *11:1*, 39–57.

Brioschi, Francesco, Luigi Buzzacchi, and Massimo M. Colombo, "Risk Capital Financing and the Separation of Ownership and Control in Business Groups," *Journal of Banking and Finance*, 1989, *13*, 747–772.

Brusco, Sandro and Fabio Castiglionesi, "Liquidity coinsurance, moral hazard, and financial contagion," *The Journal of Finance*, 2007, *62* (5), 2275–2302.

Cabrales, Antonio, Piero Gottardi, and Fernando Vega-Redondo, "Risk sharing and contagion in networks," *The Review of Financial Studies*, 2017, *30* (9), 3086–3127.

Cifuentes, Rodrigo, Gianluigi Ferrucci, and Hyun Song Shin, "Liquidity risk and contagion," *Journal of the European Economic Association*, 2005, *3* (2-3), 556–566.

Cohen-Cole, E., E. Patacchini, and Y. Zenou, "Systemic Risk and Network Formation in the Interbank Market," *mimeo: University of Stockholm*, 2012.

Covi, Giovanni, Mehmet Ziya Gorpe, and Christoffer Kok, "CoMap: Mapping Contagion in the Euro Area Banking Sector," *mimeo: ECB and IMF*, 2018.

Craig, Ben and Goetz Von Peter, "Interbank tiering and money center banks," *Journal of Financial Intermediation*, 2014, *23* (3), 322–347.

Csóka, Péter and P. Jean-Jacques Herings, "Decentralized clearing in financial networks," *Management Science*, 2018, *64(10)*, 4681–99.

Davydenko, Sergei A., Ilya A. Strebulaev, and Xiaofei Zhao, "A market-based study of the cost of default," *The Review of Financial Studies*, 2012, *25* (10), 2959–2999.

Demange, Gabrielle, "Contagion in financial networks: a threat index," *Management Science*, 2016, *64* (2), 955–70.

D'Errico, Marco and Tarik Roukny, "Compressing over-the-counter markets," *arXiv preprint arXiv:1705.07155*, 2019.

Dette, Tilman, Scott Pauls, and Daniel N Rockmore, "Robustness and contagion in the international financial network," *arXiv preprint arXiv:1104.4249*, 2011.

Diamond, Douglas W. and Philip H. Dybvig, "Bank runs, deposit insurance, and liquidity," *Journal of political economy*, 1983, *91* (3), 401–419.

Diebold, Francis X and Kamil Yılmaz, "On the network topology of variance decompositions: Measuring the connectedness of financial firms," *Journal of Econometrics*, 2014, *182* (1), 119–134.

Duarte, Fernando and Collin Jones, "Empirical network contagion for US financial institutions," *Federal Reserve Bank of New York Staff Reports*, 2017.

Duffie, Darrell and Chaojun Wang, "Efficient Contracting in Network Financial Markets," *working paper: Stanford*, 2016.

_ **and Haoxiang Zhu**, "Does a central clearing counterparty reduce counterparty risk?," *The Review of Asset Pricing Studies*, 2011, *1* (1), 74–95.

Eisenberg, Larry and Thomas H. Noe, "Systemic Risk in Financial Systems," *Management Science*, 2001, *47*, 236–249.

Elliott, Matthew, Benjamin Golub, and Matthew O. Jackson, "Financial Networks and Contagion," *American Economic Review*, 2014, *104(10)*, 3115–3153.

_ **, Co-Pierre Georg, and Jonathon Hazell**, "Systemic Risk-Shifting in Financial Networks," *mimeo: Cambridge University*, 2018.

Erol, Selman, "Network hazard and bailouts," *Available at SSRN 3034406*, 2019.

— **and Rakesh Vohra**, "Network formation and systemic risk," *mimeo: UPenn*, 2018.

Farboodi, Maryam, "Intermediation and voluntary exposure to counterparty risk," *mimeo:*, 2017.

Fedenia, Mark, James E. Hodder, and Alexander J. Triantis, "Cross-Holdings: Estimation Issues, Biases, and Distortions," *The Review of Financial Studies*, 1994, *7* (1), 61–96.

Ferrara, Gerardo, Sam Langfield, Zijun Liu, and Tomohiro Ota, "Systemic illiquidity in the interbank network," *Staff Working Paper: Bank of England*, 2017.

Fischer, Tom, "No-Arbitrage Pricing Under Systemic Risk: Accounting for Cross-Ownership," *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 2014, *24* (1), 97–124.

Gai, P., A. Haldane, and S. Kapadia, "Complexity, Concentration and Contagion," *Journal of Monetary Economics*, 2011, *58:5*, 453–470.

Gai, Prasanna and Sujit Kapadia, "Contagion in financial networks," *Proceedings of the Royal Society A*, 2010, *466*, 2401–2423.

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

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

Gofman, Michael, "Efficiency and stability of a financial architecture with too-interconnected-to-fail institutions," *Journal of Financial Economics*, 2017, *124* (1), 113–146.

Gourieroux, Christian, J-C Héam, and Alain Monfort, "Bilateral exposures and systemic solvency risk," *Canadian Journal of Economics/Revue canadienne d'économique*, 2012, *45* (4), 1273–1309.

Greenwood, Robin, Augustin Landier, and David Thesmar, "Vulnerable Banks," 2012. NBER Working Paper No. 18537.

Hadar, Josef and William R. Russel, "Stochastic Dominance and Diversification," *Journal of Economic Theory*, 1971, *3*, 288–305.

Hirshleifer, David and Siew Hong Teoh, "Systemic risk, coordination failures, and preparedness externalities: Applications to tax and accounting policy," *Journal of Financial Economic Policy*, 2009, *1* (2), 128–142.

Ibragimov, Rustam, Dwight Jaffee, and Johan Walden, "Diversification Disasters," *Journal of Financial Economics*, 2011, *99* (2), 333–348.

Jackson, Matthew O., *The Human Network: How Your Social Position Determines Your Power, Beliefs and Behaviors*, Pantheon Books: New York, 2019.

_ **and Agathe Pernoud**, "Understanding Financial Networks and Systemic Risk," *mimeo: Stanford University*, 2019.

James, Christopher, "The Losses Realized in Bank Failures," *The Journal of Finance*, 1991, *46* (4), 1223–1242.

Jensen, Michael C. and William H Meckling, "Theory of the firm: Managerial behavior, agency costs and ownership structure," *Journal of financial economics*, 1976, *3* (4), 305–360.

Johnson, Donald B, "Finding all the elementary circuits of a directed graph," *SIAM Journal on Computing*, 1975, *4* (1), 77–84.

Kanik, Zafer, "Can Rescues by Banks Replace Costly Bailouts in Financial Networks?," *mimeo: MIT*, 2018.

Kiyotaki, Nobuhiro and John Moore, "Credit Cycles," *Journal of Political Economy*, 1997, *105(2)*, 211 – 248.

Klasing, Mariko J and Petros Milionis, "Quantifying the evolution of world trade, 1870–1949," *Journal of International Economics*, 2014, *92* (1), 185–197.

Leitner, Yaron, "Financial Networks: Contagion, Commitment, and Private Sector Bailouts," *Journal of Finance*, 2005, *60(6)*, 2925–2953.

Lorenz, Jan, Stefano Battiston, and Frank Schweitzer, "Systemic Risk in a Unifying Framework for Cascading Processes on Networks," *European Physical Journal B*, DOI: 10.1140/epjb/e2009-00347-4 2009, *71*, 441–460.

Rochet, Jean-Charles and Jean Tirole, "Interbank Lending and Systemic Risk," *Journal of Money, Credit and Banking*, 1996, *28:4*, 733–762.

Roukny, Tarik, Stefano Battiston, and Joseph E Stiglitz, "Interconnectedness as a source of uncertainty in systemic risk," *Journal of Financial Stability*, 2018, *35*, 93–106.

Soramäki, Kimmo, Morten L. Bech, Jeffrey Arnold, Robert J. Glass, and Walter E. Beyeler, "The topology of interbank payment flows," *Physica A*, 2007, *379*, 317–333.

Suzuki, Teruyoshi, "Valuing corporate debt: the effect of cross-holdings of stock and debt," *Journal of the Operations Research Society of Japan*, 2002, *45* (2), 123–144.

Upper, Christian, "Simulation Methods to Assess the Danger of Contagion in Interbank Markets," *Journal of Financial Stability*, 2011, *7* (3), 111–125.

_ **and Andreas Worms**, "Estimating Bilateral Exposures in the German Interbank Market: Is there a Danger of Contagion?," *European Economic Review*, 2004, *48* (4), 827–849.

Wagner, Wolf, "Diversification at Financial Institutions and Systemic Crises," *Journal of Financial Intermediation*, 2010, *19* (3), 373–386.

Wang, Chaojun, "Core-periphery trading networks," *Dissertation, Stanford University*, 2017.

Appendix A General Contracts Between Financial Organizations

We here discuss bank values when contracts are not restricted to debt and equity. In full generality a contract between organizations i and j is denoted by $f_{ij}(\mathbf{V}, \mathbf{p})$ and can not only depend on the value of organization j, but also on the value of other organizations. This represents some stream of payments that j owes to i in exchange for some good, payment, or investment that has been given from i to j.

A.1 Values with more General Contracting

In the case in which contracts are not restricted to equity and debt holdings, the value V_i of an organization i is then

$$V_i = \sum_k q_{ik} p_k + \sum_j f_{ij}(\mathbf{V}, \mathbf{p}) - \left[\sum_j f_{ji}(\mathbf{V}, \mathbf{p}) - S_{ji}(\mathbf{V}) V_i^+ \right] - b_i(\mathbf{V}, \mathbf{p}), \tag{8}$$

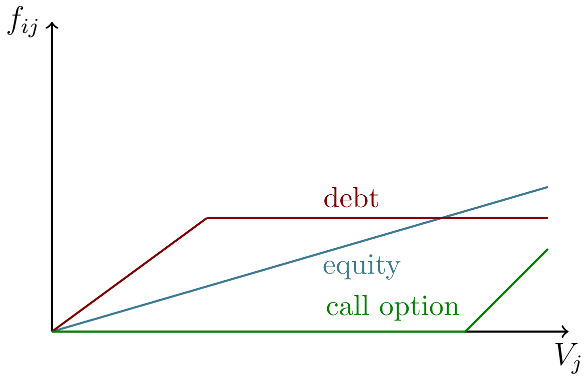
where $f_{ji}(\mathbf{V}, \mathbf{p}) - S_{ji}(\mathbf{V}) V_i^+$ accounts for the fact that debt and contracts other than equity are included as liabilities in a book value calculation.[42]

If $b_i(\mathbf{V}, \mathbf{p})$s are nonincreasing in $\mathbf{V}$ and bounded (supposing that the costs cannot exceed some total level), then if each $f_{ij}(\mathbf{V}, \mathbf{p})$ is a nondecreasing function of $\mathbf{V}$, $\sum_j f_{ji}(\mathbf{V}, \mathbf{p}) - S_{ji}(\mathbf{V}) V_i$ is nonincreasing in V, and either f is bounded or possible values of V are bounded, (using the usual Euclidean partial order) then again there exists a fixed point by Tarski's fixed point theorem for each $\mathbf{p}$. They again comprise a complete lattice. Discontinuities, which come from bankruptcy costs and potentially the financial contracts themselves, can lead to multiple solutions for organizations' values.

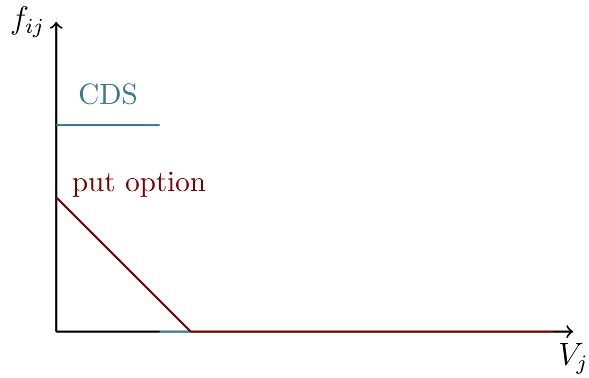
When financial contracts are not increasing functions of the values of organizations, $\mathbf{V}$, there may not exist an solution for the values of the V_is. For instance, as soon as some banks insure themselves against the default of a counterparty or bet on the failure of another, simple accounting rules may not yield consistent values for all organizations in the financial network. We illustrate this in the following example.

Example of Non-Existence of a Solution for V: Credit Default Swaps. Consider a financial network composed of $n = 3$ organizations, each of which owns a proprietary asset $\mathbf{Q} = \mathbf{I}_3$. For simplicity all assets $k \in \{1, 2, 3\}$ have the same value $p_k = 2$. The values of

[42] This more general model also embeds that of Barucca, Bardoscia, Caccioli, D'Errico, Visentin, Battiston and Caldarelli (2016) in which banks hold debt on each other, but these debt claims are not valued under full information: they allow for uncertainty regarding banks' external assets and ability to honor their interbank liabilities, whose face value may then be discounted depending on available information. Financial contracts as defined here can capture this kind of uncertainty if f_{ij} equals the expected payment from j to i given some information—e.g. a subset of known bank values or primitive asset values.



(a) Non-Decreasing Financial Contracts



(b) Non-Increasing Financial Contracts

organizations are linked to each other through the following financial contracts: organization 2 holds debt from 1 with face value $D_{21} = 1$; 2 is fully insured against 1's default through a CDS with organization 3 in exchange of payment $r = 0.4$; finally 1 holds a contract with 3 that is linearly decreasing in 3's value. Suppose an organization defaults if and only if its book value falls below its interbank liabilities, in which case it incurs a cost $\beta = 0.1$. Formally, the contracts are

$$\begin{aligned} f_{21}(\mathbf{V}) &= D_{21}\mathbb{1}_{V_1 \geq 0} \\ f_{23}(\mathbf{V}) &= D_{21}\mathbb{1}_{V_1 < 0} \\ f_{32}(\mathbf{V}) &= r\mathbb{1}_{V_1 \geq 0} \\ f_{13}(\mathbf{V}) &= -0.5V_3. \end{aligned}$$

Note that organization 2 and 3 never default: the former's value is always at least $2 - r > 0$ and the latter 's is at least $2 - D_{21} > 0$. We then check that there is no solution in which organization 1 is solvent. In such a case, $V_3 = 2 + r$ and $V_1 = 2 - 0.5V_3 - D_{21} = -0.2 < 0$: but then bank 1 defaults, which is a contradiction. Finally suppose that 1 defaults. Then $V_3 = 2 - D_{21}$ and $V_1 = 2 - 0.5V_3 - \beta = 1.4 > 0$, another contradiction.

Appendix B Proofs

Proof of Proposition 1: Let μ be the measure on $\mathbf{p}$, the vector of all portfolio values, and let

$$A(q_i) = \{\mathbf{p} \mid q_i p_i + (1 - q_i)(1 + r) + \sum_{j \neq i} S_{ij} V_j(\mathbf{p}, q_i)^+ + D_i^A(\mathbf{p}, q_i) > D_i^L\}.$$

Note that $\mu(A(1)) > 0$ by Chebychev's inequality since $\mathbf{p}$ is bounded and $\mathbb{E}[p_i] > D_i^L$ and all other variables are nonnegative. This implies that $\mu(A(q_i)) > 0$ for any possible optimizing level of q_i.

Consider any $q_i < 1$ for which $\mu(A(q_i)) > 0$ (which are the only possible optimizers), and let us examine the gain in utility that results from increasing q_i to $q_i + \varepsilon$. We show that for any such q_i there is an $\varepsilon > 0$ for which there is a gain in the expected value, and this then implies that the optimizer is 1.

Note that

$$\int_{A(q_i+\varepsilon)} V_i(\mathbf{p}, q_i+\varepsilon) d\mu(p) - \int_{A(q_i)} V_i(\mathbf{p}, q_i) d\mu(\mathbf{p}) \geq \int_{A(q_i+\varepsilon)\cap A(q_i)} \left[V_i(\mathbf{p}, q_i+\varepsilon) - V_i(\mathbf{p}, q_i)\right] d\mu(\mathbf{p}) - \int_{A(q_i)\setminus A(q_i+\varepsilon)} V_i(\mathbf{p}, q_i) d\mu(\mathbf{p}).$$

Next, note that for $\mathbf{p} \in A(q_i+\varepsilon) \cap A(q_i)$, since i is at least as dependent upon its own portfolio as others, $D_i^A(\mathbf{p}, q_i+\varepsilon) = D_i^A(\mathbf{p}, q_i)$ and

$$\sum_{j\neq i} S_{ij} \left(V_j(\mathbf{p}, q_i+\varepsilon)^+ - V_j(\mathbf{p}, q_i)^+\right) = c\left(q_i p_i + (1-q_i)(1+r)\right)$$

for some $c \geq 0$ (which follows since the V_j's depend on q_i only via linear functions of V_i).

Also, for $\mathbf{p} \in A(q_i) \setminus A(q_i+\varepsilon)$, it must be that $V_i(\mathbf{p}, q_i+\varepsilon') = 0$ for some $\varepsilon' < \varepsilon$ and that $V_i(\mathbf{p}, q_i+\varepsilon'') > 0$, for all $\varepsilon'' \in [0, \varepsilon')$. Thus, for all $\mathbf{p} \in A(q_i) \setminus A(q_i+\varepsilon)$

$$V_i(\mathbf{p}, q_i) \leq (1+c)\varepsilon'(1+r) \leq (1+c)\varepsilon(1+r).$$

Then

$$\int_{A(q_i+\varepsilon)} V_i(\mathbf{p}, q_i+\varepsilon) d\mu(p) - \int_{A(q_i)} V_i(\mathbf{p}, q_i) d\mu(p) \geq \int_{A(q_i+\varepsilon)\cap A(q_i)} (1+c)\varepsilon\left(p_i - (1+r)\right) d\mu(p) - (1+c)\varepsilon(1+r)\int_{A(q_i)\setminus A(q_i+\varepsilon)} d\mu(p).$$

Claim **1.** *If $\mu(A(q_i)) > 0$, then as $\varepsilon \longrightarrow 0$ $\mu(A(q_i) \setminus A(q_i+\varepsilon)) \longrightarrow 0$ while $\mu(A(q_i) \cap A(q_i+\varepsilon)) \longrightarrow \mu(A(q_i))$.*

Proof of Claim 1. Since each bank is at least as dependent upon its own portfolio as others, a marginal change in q_i can only make another default if it also induces i to default. Hence for ε small enough, the only bank that defaults for $\mathbf{p} \in A(q_i) \setminus A(q_i+\varepsilon)$ is bank i. This implies that for all $\mathbf{p} \in A(q_i) \setminus A(q_i+\varepsilon)$ bank i gets the same debt payments $D_i^A(\mathbf{p})$, and gets a fraction c_{ij} of other banks' portfolio value and net debt through equity claims. This fraction c_{ij} is independent of ε for ε small enough since the set of defaulting banks besides i

is unchanged. Thus we can write

$$A(q_i) \setminus A(q_i+\varepsilon) = \{\mathbf{p} : D_i^L - q_i p_i - (1-q_i)(1+r) - D_i^A(\mathbf{p}) - \sum_j c_{ij}[D_j^A(\mathbf{p}) - D_j^L] \leq \sum_j c_{ij} p_j$$

$$\text{and } \sum_j c_{ij} p_j < D_i^L - (q_i+\varepsilon)p_i - (1-q_i-\varepsilon)(1+r) - D_i^A(\mathbf{p}) - \sum_j c_{ij}[D_j^A(\mathbf{p}) - D_j^L]\}$$

where the RHS of the second inequality converges to the LHS of the first inequality as $\varepsilon \longrightarrow 0$. Hence $A(q_i) \setminus A(q_i+\varepsilon)$ converges to the set of $\mathbf{p}$ such that $q_i p_i + \sum_j c_{ij} p_j$ is equal to a constant, which has measure zero since the price vector $\mathbf{p}$ has an atomless distribution: $\mu(A(q_i) \setminus A(q_i+\varepsilon)) \longrightarrow 0$.

We now show that $\mu(A(q_i) \cap A(q_i+\varepsilon)) \longrightarrow \mu(A(q_i))$ as $\varepsilon \longrightarrow 0$. Since bank i is still solvent for $\mathbf{p} \in A(q_i+\varepsilon)$ and is at least as dependent upon its own portfolio as others, this marginal increase in its risky investment cannot induce another's default. Hence it gets the same debt payments and linear claims on others' value. Since these linear claims are bounded, $V_i(\mathbf{p}, q_i+\varepsilon) \longrightarrow V_i(\mathbf{p}, q_i)$ as $\varepsilon \longrightarrow 0$ and $\mu(A(q_i) \cap A(q_i+\varepsilon)) \longrightarrow \mu(A(q_i))$.

□

Therefore, for any $\delta > 0$, for all small enough ε the gain in utility is at least

$$\varepsilon\left[[\mu(A(q_i)) - \delta]\mathbb{E}\left[p_i - (1+r)\,|A(q_i)\right] - [\mu(A(q_i) \setminus A(q+\varepsilon))](1+r)\right],$$

which is at least

$$\varepsilon[\mu(A(q_i)) - \delta]\mathbb{E}\left[p_i - (1+r)\right] - \varepsilon(1+r)\mu(A(q_i) \setminus A(q_i+\varepsilon)),$$

which is strictly positive for small enough δ and ε, establishing the result. ▮

Proof of Proposition 2: We first show by contradiction that there cannot be an equilibrium in which banks choose independent portfolios. Suppose such equilibrium exists and consider the problem faced by bank i. Independent portfolios require the existence of at least one state of the world in which $p_i = R_i$ but $p_j = 0$ for all $j \neq i$, and similarly of at least one state in which $p_i = 0$ but $p_j = R_j$ for all $j \neq i$. We however show that bank i would be strictly better off if it were to switch its portfolio realization between two such states. Since these states are equally likely, such deviation is profitable for i as soon as

$$V_i(p_i = R_i, p_{-i} = \mathbf{R_{-i}})^+ + V_i(p_i = 0, p_{-i} = \mathbf{0})^+ > V_i(p_i = 0, p_{-i} = \mathbf{R_{-i}})^+ + V_i(p_i = R_i, p_{-i} = \mathbf{0})^+.$$

First note that

$$V_i(p_i = 0, p_{-i} = \mathbf{0})^+ = V_i(p_i = R_i, p_{-i} = \mathbf{0})^+ = 0$$

by the assumption that any bank is insolvent if all other organizations are. Thus the previous inequality becomes

$$V_i(p_i = R_i, p_{-i} = \mathbf{R_{-i}})^+ > V_i(p_i = 0, p_{-i} = \mathbf{R_{-i}})^+,$$

which is satisfied by assumption that at least one bank sees some positive value from its portfolio returns.

We now show that there exists an equilibrium with correlated assets. Given that all other banks chose correlated portfolios—i.e chose the exact same θK states in which to receive the nonzero return—we look at the problem faced by i. Note that, similarly as before, if i decides not to perfectly correlate its portfolio then there must be at least one state in which its portfolio pays off but none of the others does, and at least one state in which its portfolio does not pay off but all the others do. Hence correlation is an equilibrium as soon as the above same inequality holds *weakly* for all banks. By assumption, no bank can remain solvent if it is the only one with a positive portfolio realization, hence $V_i(p_i = R_i, p_{-i} = \mathbf{0}) = 0$ for all i. Again, the incentive condition boils down to

$$V_i(p_i = R_i, p_{-i} = \mathbf{R_{-i}})^+ \geq V_i(p_i = 0, p_{-i} = \mathbf{R_{-i}})^+,$$

which is true since V is weakly increasing in $\mathbf{p}$. Hence all banks choosing perfectly correlated portfolios is an equilibrium. ∎

Proof of Proposition 3: We first show that, unless portfolios are perfectly correlated across banks, there always exists a bank i that gets $p_i = R_i$ when others get p_{-i}, and $p_i = 0$ when others get p'_{-i} such that $|\{j \neq i : p'_j = R_j\}| > |\{j \neq i : p_j = R_j\}|$. In words, we first show that there always exists a bank i that could move one of her high portfolio realizations to a state in which there are more other banks with high portfolio realizations. Then the sufficient condition in the proposition implies directly that doing so is a profitable deviation.

Let $\Omega = \{\omega_k\}_{k=1}^K$ be the set of K states, and p_i^k the portfolio realization of bank i in state k. Let $k^* \in \arg\max_k \sum_i \mathbb{1}\{p_i^k = R_i\}$ be one of the states with highest number of high portfolio realizations. There are two cases:

(i) $k^* < n$, then there is a bank i that gets zero in state k^*: $p_i^{k^*} = 0$. Such bank could move one of her high portfolio realizations from any other state k in which $\sum_{j\neq i} \mathbb{1}\{p_j^k = R_j\} < k^*$ (by definition of k^*) to state k^*. Given the condition in the proposition, this is a profitable deviation.

(ii) $k^* = n$, then ignore all states that have k^* high portfolio realizations. Since portfolios are not perfectly correlated, this leaves at least 2 states k with $0 < \sum_i \mathbb{1}\{p_i^k = R_i\} < n$, for which the reasoning in case (i) applies.

Proof of Proposition 4:

The characterization of solvency in the best equilibrium derives directly from the definition of the best equilibrium, and the algorithm used to compute it. Recall that the first step of this algorithm is to compute each bank's value assuming all the others remain solvent and pay back their debt. Since we are focusing on networks without equity cross-holdings, these are $V_i = p_i - D_i^L + D_i^A \geq 0$ for all i if and only if the network is weakly portfolio-balanced. If a bank does not have a weakly balanced portfolio, then it must be defaulting in the best equilibrium.

The characterization of solvency in the worst equilibrium requires more work. First, since it cannot be that some banks default in the best equilibrium but not in the worst, weak balancedness of the network is also a necessary condition to have full solvency in the worst equilibrium. It is however no longer sufficient, except in the special case of a network that involves no cycles. Indeed in such a case the network is solely composed of disjoint strings, and weak balancedness is both necessary and sufficient for full solvency: it guarantees that the first bank of each string—that only has debt liabilities but no debt assets,—is unilaterally solvent which triggers a repayment cascade down the string. Hence for the following, we suppose there is at least one cycle in the network. Finally, since it is without loss of generality to prove the claim for a connected component, we suppose the network is connected.

We first show that having an iteratively strongly solvent set that intersects each directed cycle in addition to weak portfolio-balancedness implies all organizations are solvent. Let G be the set of directed edges in the financial network such that $ij \in G$ if and only if $D_{ji} > 0$. An edge from i to j means that bank i owes some debt to bank j.

First note that N can be partitioned into three sets of banks: banks that are part of at least one cycle, banks that are part of no cycle but belong to a string that eventually points to some bank in a cycle (in-going strings)[43], and banks that are part of no cycle but belong to a string that is pointed to by an organization on a cycle (out-going strings)[44]. Following the same argument as above, weak-balancedness is enough to guarantee solvency of banks on in-going strings. Furthermore, note that banks on out-going strings do not have liabilities towards banks on cycles. Hence debt payments that still (after accounting for in-going strings) have to be made to banks on cycles can only come from banks that also lie on a cycle. Now suppose there is an iteratively strongly solvent set that intersects each directed cycle, and call it B. By definition, all banks in B must be solvent in the worst case equilibrium, which means that there is at least one solvent bank on each cycle. We prove by induction on the number of banks n that this implies all banks in the network are solvent.

It clearly holds for $n = 2$, since the assumption of at least on cycle implies a single possible network configuration. One bank being solvent means that the other gets all of its incoming debts and, by weak-balancedness, can pay the full amount out.

[43]Formally an in-going string is a set of nodes $X \subseteq N$ that can be partitioned into K elements $X = X_1 \cup X_2 \cup \cdots \cup X_K$ such that nodes in X_1 are not pointed to by anyone ($D_i^A = 0$ for $i \in X_1$) and nodes in X_k are only pointed at by nodes in $X_1 \cup X_2 \cup \cdots \cup X_{k-1}$.

[44]Formally an in-going string is a set of nodes $X \subseteq N$ that can be partitioned into K elements $X = X_1 \cup X_2 \cup \cdots \cup X_K$ such that nodes in X_K point at anyone ($D_i^L = 0$ for $i \in X_K$) and nodes in X_k only point at by nodes in $X_{k+1} \cup X_2 \cup \cdots \cup X_K$.

Now suppose the claim holds for a network of size up to $n-1$, we show it holds for $n \geq 3$. Pick any bank i_0 that *is* solvent and lies on a cycle in the network with n nodes.

Let X_1^{in} be the set of nodes that only point at i_0 and X_1^{out} the set of nodes that are only pointed at by i_0 (which could be empty.) Iteratively, define X_t^{in} to be the set of nodes that only point at members of X_{t-1}^{in}, and similarly for X_t^{out}. Since the network is finite this process must terminate. Importantly note that all nodes in $X^{in} \equiv \cup_t X_t^{in}$ only point at nodes in $X^{in} \cup i_0$, and all nodes in $X^{out} \equiv \cup_t X_t^{out}$ are only pointed at by nodes in $X^{out} \cup i_0$. Either or both of these sets could be empty.

There are two possible cases:

1. The subgraph found by removing $X^{out} \cup X^{in} \cup \{i_0\}$ contains no cycle. In that case i_0 being solvent clears the entire system: if i_0 is solvent then all banks in X_1^{out} are solvent as well. If X_t^{out} is solvent then all banks in X_{t+1}^{out} are solvent as well, and more generally all banks in X^{out} are solvent. Recall that in-going strings are always solvent by above argument. Since all banks in X^{out} are solvent, they repay their debts in full, and the first organizations in any remaining out-going string must receive their debt payments in full. By weak-balancedness, they are then solvent as well, and this cascades down the string: the system clears, and the claim holds.

2. The subgraph found by removing $X^{out} \cup X^{in} \cup \{i_0\}$ contains at least one cycle. Any isolated string that is generated by removing $X^{out} \cup X^{in} \cup \{i_0\}$ must be solvent by weak-balancedness, and argument in 1. Any other component contains at least one cycle. We claim that having i_0 as well as one bank per cycle in the remaining subnetwork is enough to ensure solvency of the original network. First, recall that i_0 solvent means that all banks in X^{out} are solvent, and hence that all debt payments from banks in the removed $X^{out} \cup X^{in} \cup \{i_0\}$ to banks in the remaining subnetwork are made in full. This ensures that the remaining subnetwork is still weakly balanced. Now by assumption, having one bank per simple cycle in this subnetwork is enough to ensure its full solvency: all debt is paid back within the subnetwork. The last banks added to X^{in} have debt coming from outside of X^{in} only, and thus they are solvent. This spreads through X^{in} and eventually reaches i_0. Hence the system clears, and the claim holds: weak-balancedness as well as having one solvent bank per cycle guarantees systemic solvency.

Finally, we prove that if there does not exist an iteratively strongly solvent set that intersects every cycle, then some banks default in the worst equilibrium. First note that the union of iteratively strongly solvent sets is also an iteratively strongly solvent set: hence there exists a maximal one which, by assumption, does not intersect every cycle and hence does not include all banks. We claim that the maximal iteratively strongly solvent set is then actually the set of solvent banks in the worst equilibrium, and that this comes directly from the definition of the algorithm used to derive the worst equilibrium. Indeed, assuming no bankruptcy costs, the algorithm to compute the worst equilibrium first assumes no debt is repaid. This entails only unilaterally solvent banks remain solvent. Iterating on this, only banks in the maximal iteratively strongly solvent set remain solvent in the worst equilibrium.

If such set is not equal to N, then some banks must be defaulting. ∎

Proof of Corollary 1: Proposition 4 gives necessary condition to have systemic solvency if the best and worst equilibrium. Hence the minimum necessary bailout needed to ensure solvency in each case are such that the resulting network satisfies these necessary conditions. For the best equilibrium, this only requires making the network weakly portfolio balanced, which means rebalancing each bank's portfolio by injecting $[D_i^L - D_i^A - p_i]^+$ in each bank i. For the worst equilibrium, it also requires enough capital to generate an iteratively strongly solvent set that intersect each directed cycle.

To prove (iii), first note that once the network has been cleared of all cycles, it is only composed of strings. Suppose the claim is not true, such that the best and worst equilibria differ. For them to differ, it has to be that at least some bank i defaults in the worst equilibrium, but not in the best. For this to be true, it has to be that $p_i + D_i^A \geq D_i^L > p_i$, and that it remains solvent in the best equilibrium because once of its debtor j is solvent in the best equilibrium and repays its debt to i, but defaults in the worst equilibrium. Iterating the same reasoning one step up the string, bank j is solvent in the best equilibrium but defaults in the worst one only if one of its own predecessor does as well. Iterating up until one reaches the beginning of the string, this implies that one of the banks at the origin of the string defaults in the worst equilibrium, but remains solvent in the best one. This is however impossible since such bank has no liability, and is either unilaterally solvent, or defaults in *every* equilibrium.

A Note on the Existence of Values Satisfying Equation (3): $(\mathbf{I} - \mathbf{S})$ is invertible if and only if the matrix power series $\sum_{k=0}^{\infty} \mathbf{S}^k$ converges, which is equivalent to the largest eigenvalue of $\mathbf{S}$ in absolute value being strictly below one. Denote by λ the largest eigenvalue, and v the associated eigenvector. From the Perron Frobenius theorem, we know that $\lambda \geq 0$ and v can be chosen to be nonnegative.

By contradiction suppose $\lambda \geq 1$. Then $\sum_k S_{ik} v_k = \lambda v_i \implies \sum_k v_k \sum_i S_{ik} \geq \sum_i v_i$. Since banks are either private or public, this is equivalent to $\sum_{i\,public} v_k \geq \sum_i v_i$. To have a contradiction, we need to show that there exists a private bank i with $v_i > 0$. Since the eigenvector v cannot be a vector of zeros, it must be that $v_k > 0$ for some k. If k is a private firm, then we get a contradiction directly. If k is public, then by assumption there must be an ownership path from a private bank i to k. The private bank must then have $v_i > 0$, which contradicts $\sum_{i\,public} v_k \geq \sum_i v_i$. Hence $\lambda < 1$ and $(\mathbf{I} - \mathbf{S})$ is invertible. ∎

Appendix C Additional Discussions

C.1 Inefficient Bank Size

Mergers can affect the interdependencies in the financial network in complex ways. Here we examine banks' incentives to merge at an ex ante stage, before a cascade. This complements an analysis by Kanik (2018) who discusses banks' incentives to merge to save themselves from failure in a cascade.

The analysis in Section 3.3.1 can be seen as a sort of merger analysis - as all banks in those syndicates have identical outcomes.

More generally, we can look at the effect of a merger between bank i and j into a larger organization k such that $D_{k\ell} = D_{i\ell} + D_{j\ell}$ and $D_{\ell k} = D_{\ell i} + D_{\ell j}$ for all $\ell \neq i, j$, and similarly for equity shares.

How bank size interacts with the fact that banks choose overly risky investments is ambiguous. Indeed a merger can affect a bank's choice of risky investment in either direction: it can lead the merged organization to invest a larger or smaller share of its portfolio in the risky asset. It has generally no effect since most organizations invest fully in the risky asset irrespective of the network structure (Proposition 1). It is however possible to find examples in which size, i.e. a merger, matters. For instance consider another version of the example in Section 4.3 with $n = 3$ banks. Bank 3 can only invest in the safe asset; it has equity share s in bank 1 and debt claim $\underline{d}$ in bank 2. It owes $\overline{d} > \underline{d}$ to bank 2 as well. Suppose $1 + r + \underline{d} < \overline{d}$ such that bank 3 defaults if its equity claim on bank 1 does not yield anything. In the only decentralized equilibrium, both bank 1 and 2 invest fully in the risky asset.

Now suppose bank 1 and 2 merge, and call this new organization bank 4. Bank 4 can always prevent the default of bank 1 by investing a minimum amount in the safe asset. If this required safe investment is small enough—i.e. if $\varepsilon := \overline{d} - \underline{d} - 1 - r$ is small enough—doing so can be optimal: bank 4 may optimally set $q_4^* < 1$, and the merger may reduce investment in the risky asset. Finally suppose bank 3 and 4 merge. Then there only remains a single bank, whose optimal portfolio must have all its capital invested in the risky asset. So in general a merger can change incentives in either direction.

Mergers can however mitigate some of the inefficiencies coming from over-correlation of bank portfolios (highlighted in section 4.3). Indeed the incentive to correlate investments straight-forwardly disappears if the two banks merge: the larger organization then only invests in the asset with highest expected return.

Finally size of banks matters when analyzing contagion: larger banks can serve as buffers and stop default cascades, or on the contrary be brought to insolvency by one of their smaller branches. If the two merging banks have debt claims on each other then the merger also decreases their insolvency threshold, which reduces the likelihood of default all else equal. A merger between bank i and j increases the set of defaulting banks if and only if one of the two banks—say bank i—would have remained solvent had the merger not happened whereas they both default once merged. This requires $X_i - D_{ji} \geq 0$ but $X_j + D_{ji} - D_{ij} < 0$, where X_i is the net value of i's asset excluding its debt contracts with j, and similarly for X_j. Bank j brings i to insolvency during the merger if

$$X_i + X_j < 0 \implies X_j < -D_{ji}$$

that is, if the net debt that j owes to other banks is at least as large as what it used to owe to bank j. Hence in general, a merger can either mitigate or worsen a default cascade.