What Makes Financial Markets Special? Systemic Risk and its Measurement in Financial Networks

Matthew O. Jackson and Agathe Pernoud*

February 2019

Abstract

We provide an overview of some key trends and features of financial networks related to systemic risk. We also provide a new network model of inter-dependencies in which financial institutions choose the risk in their portfolios, their trading partners, and the correlation of their portfolios with those of their partners. We show that they have incentives to choose too much risk and correlation, and to under diversify in terms of trading partners. We also provide a measure of financial centrality in terms of the consequences of a given institution's portfolio on the potential defaults of other institutions. We discuss minimum interventions or capital needed to ensure systemic solvency.

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

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

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 (as detailed below). It is not a coincidence that the world poverty rate has fallen

*Department of Economics, Stanford University, Stanford, California 94305-6072 USA. Jackson is also an external faculty member of the Santa Fe Institute and a fellow of CIFAR. Email: jacksonm@stanford.edu and agathep@stanford.edu. We gratefully acknowledge financial support under NSF grant SES-1629446.

[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.

from over 40 percent in the early 1980s to below 10 percent today.[2] This trend has been matched by a tenfold decrease in the incidence of wars, by multiple measures - as trading partners rarely go to war with each other.[3]

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, led to a broad financial crisis and prolonged recession.[4]

Over these same decades there has been a growth in the study of networks, including studies of economic and financial networks.[5] These have provided us with the tools to understand and to quantify systemic risk. In this paper, we provide an overview and synthesis of some of the main insights that have emerged about financial networks and systemic risk. In doing so, we also provide a new model and results on incentives in financial markets.

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 consequence. At a most basic level, the risk that a counterparty defaults has consequences in a world without complete markets in which many risks cannot be hedged.[6,7] Defaults involve substantial inefficiencies and costs (Section 2.2), and can lead to potential disasters if left to cascade. These introduce externalities, since if one organization has poor judgment in its investments, poorly managed business practices, or even just unusually bad luck, this ends up affecting the values of its partners, and their partners, in discontinuous ways. Costs of avoiding cascades can be much less if addressed before they begin. However, this involves having a detailed view of the network of financial inter-dependencies, 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. These are the focus of what follows.

[2]World Bank Poverty Report.

[3]See Jackson and Nei (2015) for more empirical background and analysis of the relationship between increased trade and decreased interstate armed conflict.

[4]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).

[5]For detailed overviews of the broader literature see Jackson (2008, 2019). Specific references on financial networks are provided throughout this paper. What follows complements the discussion in Chapter 4 of Jackson (2019), which details a financial crisis and discusses some key aspects of financial markets and policy prescriptions. Some of the same points are covered here, but at a more technical level and with a model that explores several new aspects of incentives in financial markets.

[6]As a poignant example, there are even risks that the insurance and hedges meant to help eliminate risk will fail. For instance, a key failure in the financial crisis in 2008 was that AIG had sold insurance on many contracts that it was unable to deliver. 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 equilibrium existence, as we discuss in Section 3.6.

Below we begin by providing a basic overview of globalization and the growth of the financial network, in terms of connections between countries, between financial institutions themselves, as well as in terms of the size of financial institutions, the concentration of capital among the largest ones, and the patterns of their interactions.

We then provide a brief discussion of some of the central lessons from the literature on financial networks to date, especially with respect to how network structure matters. Here we highlight which aspects of financial network are similar to other network contagion settings, as well as how they are fundamentally different. In basic contagion settings, adding more connections increases the opportunities for spread and diffusion, and leads to higher rates of contamination. In financial networks, there are countervailing effects: increasing the number of counterparties diversifies an organization's portfolio. We also discuss how contagion depends on the network topology, the extent of exposure, and the type of shock that hits the system. In addition, we distinguish different ways in which there can be contagion in a market: not only in terms of cascading defaults, but in terms of lost confidence in the market or in others' faith in the market; as well as in inferences made across similar institutions.

We then introduce a model of financial networks that involve two different forms of contracts: debt and equity. This generalizes existing models in a way that is important, since we then explore how debt and equity lead to different incentives for investment as well as different probabilities of cascades, for the same initial conditions. We examine three basic choices of banks or other financial institutions: which investments they make, how many counterparties they transact with, and the extent to which they choose the same portfolios as their counterparties. We show how externalities result in inefficiencies in all of these choices, and discuss the consequences in light of the network.

We close the paper with a discussion of the lessons from the model. We discuss how the model can be used to measure systemic risk from a potential shock or default. We also discuss what minimum costs are to a government that needs to bail out an insolvent network. Here we emphasize that the methods suggested are actually straightforward, but that they require detailed information about the full network and that having partial balance sheets from firms – without the precise listing of all counterparties – is insufficient to properly manage the network. It is impossible to approximate systemic risks without detailed network information.

2 Background on Systemic Risk

We begin with some background on global trends and then discuss some of the key insights from the literature on systemic risk in financial markets.

2.1 Some Background on Globalization and Financial Interconnectedness

2.1.1 Globalization and Financial Interdependencies

As mentioned in the introduction, the growth in international trade has been shadowed by a similar growth in international finance. For example, 17 percent of equities and 18 percent of bonds around the world were held by foreigners in 2000, and that rose to 27 percent of equities and 31 percent of bonds in 2016. This matches up with the amount of investments (debt, equity, FDI, lend/other) around the world that comes from foreign sources; which was 26 trillion dollars in 2000, grew to 102 trillion dollars by 2007, and reached over 132 trillion dollars in 2016, out of a total level of world investments of just over 300 trillion dollars.[8] Thus, a rough rule of thumb is that about a third of a country's finance is now from foreign sources.

In addition, the financial sector is characterized by strong interdependencies - so capital is 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*.[9] As another example, in Germany domestic inter-bank positions account for 13% of the banking system's total assets (Craig and Von Peter (2014)). Furthermore, more than three quarters of German banks both lend to and borrow from other banks, suggesting that these interdependencies are complex and multidimensional.

2.1.2 Consolidation

The financial/banking sector has grown enormously, but has also consolidated, with far fewer banks and those being much larger than they used to be. In 1980 there were 14 thousand commercial banks in the US according to the FDIC[10], with total assets of 2 trillion dollars. In 2018 there were 4.7K with 16.5 trillion dollars in assets. So, the number of banks has dropped to a third of the previous number and at the same time are managing more than eight times as much in terms of total assets.[11] This consolidation has continued to grow even after the 2008 financial crisis. For example, in 1990 the five largest banks in the US had 10 percent of the assets, in 2007 they had 35 percent, and in 2015 45 percent. The ten largest banks in the world controlled 26 trillion dollars in 2016. To put that in perspective,

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

[9] The large difference reflects the fact that many other types of financial institutions 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.

[10] See https://www5.fdic.gov/hsob/HSOBRpt.asp

[11] Part of this change is due to changes in regulations, such as undoing the separation of investment, commercial banking, and insurance, that had been required under Glass Steagall. However, this trend is also seen outside of the US, reflecting large economies of scope and scale in the banking sector.

the US and Chinese combined GDP in 2016 was 29 trillion dollars, and the world GDP was 75 trillion dollars.

2.1.3 Core-Periphery Structure

Empirical studies of financial networks, and especially inter-bank lending, have highlighted the core-periphery structure in which most of financial trade is realized by a few large banks. For instance, Soramäki et al. (2007) detail a completely connected core of 25 banks, including all of the largest ones, that borrow and lend with each other; with large exposures between the largest banks. These core banks are highly interconnected whereas the rest of the network is usually sparse, with smaller regional banks each only interacting with a few of the larger core banks (see for instance Bech and Atalay (2010) for the US, Craig and Von Peter (2014) for Germany, and Blasques et al. (2018) for the Netherlands.)

There are a variety of good reasons to have a core-periphery structure as there are advantages to having a concentration in intermediaries as they can better manage their inventory and match buyers with sellers (e.g., see Craig and Von Peter (2014); Babus and Hu (2017); Farboodi (2014); Wang (2017)).[12] However, the dense connections within a core-periphery structure also have contagion consequences. For instance, Elliott, Golub and Jackson (2014) show how the core can lead to much more extensive default cascades for wider ranges of transactions than more balanced networks. This is exacerbated by the fact that core organizations often have very similar businesses and thus very correlated investments, which leads them to be vulnerable at the same time, as we discuss in more detail below.

2.2 Bankruptcy Costs

The externalities matter not only because of the basic investment distortions that result, but also because of the substantial frictions and costs of bankruptcy that are present in financial networks. If a large counterparty of some financial organization falls into bankruptcy, that can result in large losses for the organization, and ultimately cause it to default on some payments as well.

As an example, in Lehman Brothers bankruptcy there were initially 1.2 trillion dollars of claims made. Of these, the courts ultimately allowed only 362 billion dollars of claims and creditors only received 28 percent of that reduced number.[13] This was an extreme case, but there are substantial frictions, delays, and inefficiencies that result from bankruptcy, especially in troubled times. These result from fire sales, early termination of contracts, the complexity of contracts that need to be unwound, lengthy negotiations, legal costs, among others.

[12] Wang provides extensive references to studies documenting the presence of core-periphery structures in wide varieties of over-the-counter markets.

[13] Fleming and Sarkar (2014)

Estimates of bankruptcy recovery rates are in the 56-57 percent range.[14] Of the amounts that are lost about 4/11 is attributable to legal costs and the other 7/11 to a drop in asset value (some from liquidation). Moreover, recovery rates are another 15 to 22 percent lower in distressed times,[15] which would typically apply to a large financial crisis.

2.3 Three Types of Systemic Risk

Financial markets can exhibit at least three broad classes of systemic failure, in each of which externalities are central.

The first is the classic form of bank runs and panics, in which behavior becomes self-fulfilling. Classic treatments of this range from Keynes (1936) to Diamond and Dybvig (1983).[16] As we shall see below, in many financial settings there can exist multiple equilibria, and people's expectations about how assets will be valued can come to be self-fulfilling.[17] In classic examples, people have to forecast how others will behave, as for instance if they believe that others will all withdraw their money and leave a bank insolvent, then they have an incentive to do so as well.

A second form of contagion relates to uncertainty, where information about one institution being insolvent leads to a form of "guilt by similarity" - in which one then has doubts about the solvency of other enterprises that are somewhat similar, whether or not that inference is justified.[18] One also has a fear that the lending between institutions might freeze up, and so one begins to hoard whatever cash and assets one has, which leads to further tightening and potential spiraling. This was present in the freezing up of overnight lending during the period from 2007-2009 (e.g., see the discussion in Brunnermeier (2009); Diamond and Rajan (2011).) Not only did lending dry up, but many stock markets around the world lost nearly half or more (e.g, in the case of the Dow) of their value, while the underlying fundamentals did not reflect such a dramatic drop.

The most basic form of contagion, is not built on expectations or inference, but instead on actual cascades of insolvencies. A bank becomes insolvent and cannot pay its debts. As those liabilities are defaulted upon, this worsens the balance sheets of other institutions leading some of them to become insolvent. As more of them become insolvent, this further worsens conditions for others and the situation can potentially become catastrophic. This can happen not only because of direct defaults, but also because of forced sales of assets, which depress the values of others' portfolios. As inabilities to make payments cascade, large amounts of government intervention can be needed to avoid systemic failure, as for instance

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

[15] Acharya, Bharath and Srinivasan (2007); Bruche and Gonzalez-Aguado (2010).

[16] For more background see Reinhart and Rogoff (2009).

[17] In some circumstances, one can refine the uncertainty and produce unique predictions of self-fulfilling runs, as in Morris and Shin (1998).

[18] For instance, see Caballero and Simsek (2013) and Alvarez and Barlevy (2015). For more general background on co-movement of firms' values as well as network positions, see Diebold and Yılmaz (2014).

with the interventions in AIG, FNMA, and FHLMC in the 2008 financial crisis.

Of course, all three of these forms of contagion interact and are often at play at the same time, but the third form is one in which the actual network topology of counterparty relationships matters. Our focus here is mainly on the network of interactions and the structural implications for systemic risk, but we will discuss all three of these at different points in what follows.

2.4 Network Patterns and Systemic Risk

The literature on how systemic risk relates to network patterns has grown in recent years.[19]

There are a few points that are echoed throughout, and worth emphasizing since they are important background for the analysis that follows.

The most basic and obvious one is that the network structure matters. The more interesting part is that there is a nonmonotonicity in the potential for contagion in the density of the network.[20] This is one thing that is special about financial markets and distinguishes them from more basic analyses of contagion and diffusion, for instance of disease, in which adding more interactions only leads to more extensive rates of infection. As a bank adds more counterparties there are countervailing effects. It does become more susceptible to shocks from more sources - and hence this tends to increase the potential for cascades. However, it also becomes less exposed to any given counterparty, and so this lowers the potential for contagion.

We illustrate this in the context of a very simple example that illustrates a few main points. Consider a network of identical banks that have balance sheets of the form in Figure 1.

On the liability side, each bank has 10 units of capital from deposits and another 10 units of capital from loans from other banks and their owners have 2 units of capital in the form of equity. On the asset side, the banks each have an investment portfolio worth 12 and loans to other banks worth 10.

Now let us suppose that the investment portfolio of some bank drops in value and ends up worth less then 10, as for instance in Figure 2. This bank is now insolvent, as its assets are worth less than its liabilities, and so it defaults on some of its payments. For the purposes of

[19] 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), Cabrales, Gottardi, and Vega-Redondo (2017), Kanik (2018).

[20] For variations on this theme see Allen and Gale (2000), Cifuentes, Ferrucci and Shin (2005), Gai and Kapadia (2010), Elliott, Golub, and Jackson (2014), Cabrales, Gottardi, and Vega-Redondo (2017), and Gofman (2017).

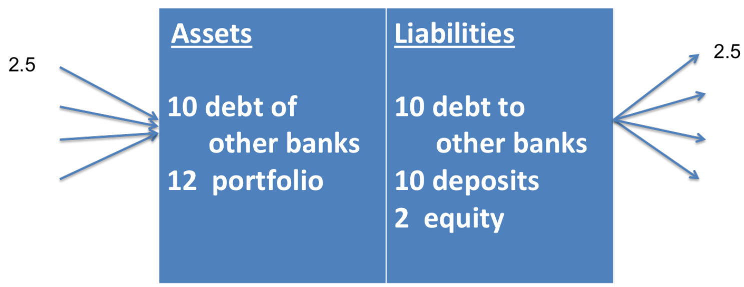


Figure 1: The starting balance sheet of the banks in a network.

this example, let us treat the default as total on at least one of its loans, although one can obviously rescale the numbers and extend the example to work with some partial default.

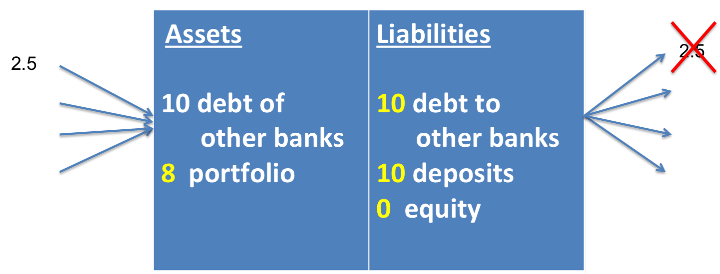


Figure 2: Some bank's portfolio drops to a value, below 10, say to 8. This makes the bank insolvent, and so it defaults on some of its payments.

Initially the bank owed four different banks 2.5, and so it fails to make one of the 2.5 payments. This then has to be written off by the second bank that made the loan to the first bank, and so that second bank loses 2.5 in terms of its assets. The second bank is now insolvent as well, as the value of its assets are only 19.5 and it has 20 units of liabilities. So, this second bank now defaults on some payments, as pictured in Figure 3. This now cascades.

With the exposure of 2.5 to each other bank, and only an initial equity value of 2, banks are susceptible to even a single defaulting counterparty. As more banks are hit by the defaults of other banks, this cascades.

There are two key elements that lead to the cascading defaults in this example. One is that banks have nontrivial exposures to other banks relative to the size of their portfolios. The second is that they each have several counterparties, but also not too many counterparties. The fact that they owe 10 to four counterparties makes the value of each loan 2.5, which is beyond their equity value, so even one defaulting counterparty is enough to make them insolvent. Also, the fact that they each owe four other banks 2.5, means that this cascade can grow in size as each defaulting bank infects others.

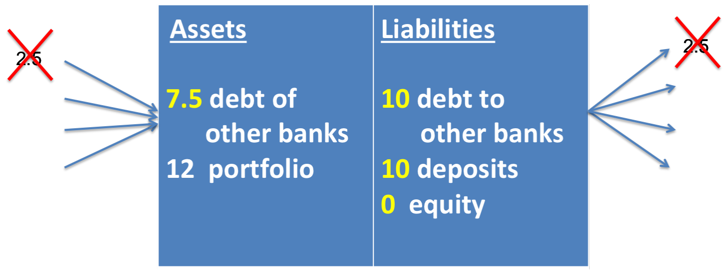


Figure 3: A second bank now becomes insolvent due to its lost asset value from the loan from the first bank. It then defaults on some of its payments.

To see the importance of the limited number of counterparties, let us alter the example so that each bank has ten other banks to which it owes 1 unit each, as in Figure 4. In this case, there is no longer any cascade. The default by any single counterparty no longer leads a bank to become insolvent.

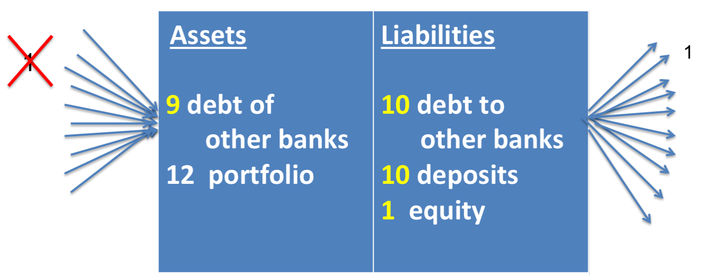


Figure 4: Even though banks have more counterparties, the lower exposure to each separate one of them now makes them immune to a default by any single counterparty.

Here we see the non-monotonicity quite clearly. We have increased the number of counterparties of each bank, and hence made the financial network denser, and yet have eliminated the cascade.

This effect distinguishes financial contagion from other sorts of contagions. With simple contagions, such as those of a flu, increasing the number of connections leads to a denser network and more opportunities for spreading, leading to faster and wider potential contagions. Even with diffusion of something that takes several interactions to lead to infection – for instance, hearing a rumor several times before believing it and passing it along – increasing the number of contacts and the overall density of a network increases the potential contagion.[21] In contrast, with financial networks the increase in the number of counterparties,

[21] See Centola (2018) and Jackson (2019) for background discussion and references on the contrast between

holding fixed the total amount of exposure, leads to lower levels of contagion for the same shock.

One important question that then arises is whether financial institutions have the right incentives to spread their trades around among many diverse counterparties, or whether they have incentives to overly concentrate them. As we shall show below, the answer to this question is in the negative.

We can also see other aspects that are important in financial networks. We could also have had four counterparties for each bank, but instead had only a total of 4 units of lending on each bank's balance sheet, so 1 unit to each of four other banks. By lowering the amount of exposure, we could have also avoided contagion. This is a point made in Elliott, Golub, and Jackson (2014). Increasing the amount of exposure of each bank to others tends to increase the propensity for contagion, as it does in this example. However, as Elliott et al. (2014) also discuss, it can help diversify a given bank's portfolio. Thus, there can be a distinction in that more exposure and trading can help diversify any given bank's portfolio and thus make their own investments less variable and more stable, but also leads to an increase in the possibility of contagion.

Along with this are results that the size of the shock matters and different networks may have different susceptibilities depending on the shock. On an intuitive level, adding more connections helps with diversification and thus lowers risk as in the above example, except if there are larger shocks, in which case the added connections simply aid in transmitting that shock, but there are further nuances that depend on the model and the type of contracts that exist between institutions.[22]

This possibility of contagion is particularly affected by the presence of a core-periphery structure, which is common in many financial settings as discussed above. Large core banks can be resistant to small shocks, but then can fail catastrophically when hit with larger shocks. As shown in Elliott, Golub, and Jackson (2014), this can erase the non-monotonicity discussed above, since the failure of a large bank or entity that other core banks have large exposure to can lead to extensive contagion within the core, and then by extension to the whole economy. This was what loomed in 2008.

Finally, let us mention how correlation in portfolios plays into the picture. The point is easy to see in the context of the above example. Correlation in portfolios can mitigate and even erase the non-monotonicity in contagion in financial networks. Let us reconsider the scenario in Figure 4. Suppose that the portfolios of banks exhibited substantial correlation – for instance, all having substantial exposure to the same sort of CDOs as was the case in 2007. For example, suppose when one bank's portfolio of investments is down and drops below 10, it is also very likely that the portfolios of other banks are at below-normal levels, which can even be exacerbated by a fire sale of the first bank. If we change the portfolio of the second bank pictured in Figure 4 to drop to 10 at the same time that the first bank's

simple and complex contagion, and see Jackson and Storms (2017) for a more nuanced look at that contrast.

[22]For instance, see Allen and Gale (2000), Gai and Kapadia (2010), and Acemoglu, Ozdaglar and Tahbaz-Salehi (2015).

portfolio dropped to 8, then the default of the first bank is enough to push the second bank into insolvency. This is pictured in Figure 5

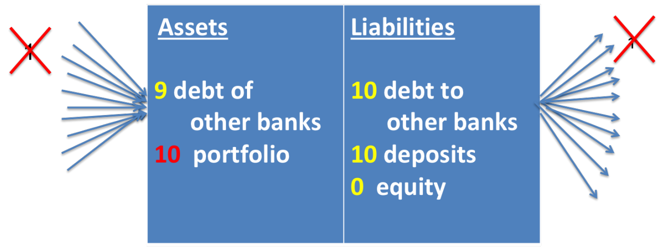


Figure 5: With correlated portfolios, banks are now more susceptible to defaults of others, even when levels of exposure to any single counterparty are low. This can undo the non-monotonicity discussed and now a second bank defaults even with ten partners.

This sort of correlation in portfolios can thus be very problematic, and is seen in the data (e.g., see Duffie, Eckner, Horel and Saita (2009); Elliott, Georg and Hazell (2018)). It occurs for many reasons. First, as we will see shortly, banks have incentives to choose portfolios that are correlated with those of their counterparties.

Second, competition between the institutions can lead them to choose similar investments. This was something that contributed to the Savings and Loan exposure to junk-bonds and market crisis of the 1980s. This happened since savings and loans that took riskier positions could offer higher interest rates on their checking and savings accounts. Since many of these were insured accounts, depositors had incentives to shop for the highest interest rate. This means that in order to attract and keep depositors, savings and loans had to compete to offer the highest interest rates. Since the place they could earn the higher expected rates of return that would enable them to offer higher interest rates was in riskier investments, and in the junk-bond market in particular, this drove them to take increasingly risky positions. This incentive is not unique to the savings and loan crisis, but is more commonly at work in the banking sector.

Third, there can be regulations of which sorts of investments that banks or other financial institutions can make with some of their capital. If banks are all required to hold some percentage of their portfolio in bonds issued by some countries (e.g., European countries) then it is not surprising that they all have nontrivial amounts of investment in Greek (or today Italian) bonds. This can correlate their portfolios and make them susceptible to the same shocks.

2.5 Externalities and Financial Contagion

Given that systemic risk depends on the network, investments of the organizations, and their correlation, it is worthwhile to understand the extent to which institutions have efficient incentives - i.e., those to make investments and to choose partnerships that maximize the overall value of the society.

Given that there are externalities here, we should expect individual financial incentives to fail to align with the overall welfare of the economy.[23] Indeed, we show that incentives are misaligned on all dimensions.

3 A Model of Financial Interdependencies

Here we introduce a model of financial interdependencies that allows us to examine the choices of financial institutions in the network. In the model we include both debt and equity since there are important distinctions in the incentives and systemic risk they generate on the network.[24] Indeed, many forms of contracts, including some swaps, can be approximated as one or the other. There are obviously more complex contracts that could be built into such a model. We describe a general form of the model for such more complex contracts as well, but most of the insights can be most crisply obtained with debt and equity.

It is also important to include both debt and equity since both types of securities are needed to capture the balance sheets of some of the most prominent and important different types of financial institutions. For example, banks' balance sheets involve substantial portions of deposits, loans, CDOs (collateralized debt obligations), and other 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.

3.1 Primitive Assets, Organizations, and Cross-Holdings

Consider a set $N = \{0, 1, \ldots, n\}$ of institutions involved in the network.

We treat $\{1, \ldots, n\}$ as the financial organizations, or "banks" for simplicity in terminology. But as mentioned above, these should be interpreted as a broad variety of financial institutions, ranging including banks, venture capital funds, broker-dealers, and many other sorts of institutions that have substantial financial exposures on both sides of their balance sheets. These are organizations that can issue as well as hold debt, and buy and sell equity and make other investments.

[23] This is a general theme in the network formation literature, starting with Jackson and Wolinsky (1996), but plays out slightly differently here. For additional discussion of the efficiency or inefficiency of network formation in financial settings in other models, see Babus (2016); Gofman (2017); Erol (2018); Erol and Vohra (2018); Kanik (2018).

[24] By having both debt and equity, the model generalizes those built with debt-like interdependencies (e.g., Eisenberg and Noe (2001), Gai and Kapadia (2010)), Csoka and Herings (2018), and those with equity-based interdependencies (e.g., Elliott, Golub and Jackson (2014)).

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).[25] 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 institutions that are the dominoes, becoming insolvent and defaulting on payments as a result of defaults on their assets.

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.) Portfolios and realized cash-flows determine each organization's total equity value V_i, whose computation will be explicitly defined below.

An organization can also have financial contracts with other organizations in the network. In general 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 or investment that has been given from i to j.

For instance if bank i owns equity share in bank j, then $f_{ij}(\mathbf{V}, \mathbf{p}) = S_{ij}V_j$ 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.

3.2 Values in a Network of Debt and Equity

For the sake of tractability, we begin by restricting the set of possible financial contracts to debt and equity. 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. Organizations are either privately owned $\sum_i S_{ij} = 0$, or public $\sum_i S_{ij} = 1$. Suppose there exists at least one private

[25] There is a complex spectrum that involves a lot of gray area. For instance, Harvard University invests billions of dollars, including making large loans. At the same time it borrows money and has issued debt. 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 and others that do not. 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 arbitrary jurisdictional rules.

organization, and that every public organization has some indirect private ownership—i.e. there exists a directed path in equity from every public bank to a private bank. A financial network is then a tuple $(N, \mathbf{D}, \mathbf{S})$. Finally denote by

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

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.

The book value V_i of an organization i 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 :

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

Equation (1) can be written in matrix notation as

$$\mathbf{V} = \mathbf{Q}\mathbf{p} + \mathbf{D}^A - \mathbf{D}^L + \mathbf{S}\mathbf{V}.$$

Solving this,[26]

$$\mathbf{V} = (\mathbf{I} - \mathbf{S})^{-1} \left[\mathbf{Q}\mathbf{p} + \mathbf{D}^A - \mathbf{D}^L\right]. \tag{2}$$

Written this way, the book value of an organization coincides with its 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

[26]To see that $(I - S)$ is invertible, and to ensure that all of the V's are bounded (which we will need to apply Tarski's theorem later on), 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 since summing them up 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.

3.3 Discontinuities in Values and Failure Costs

We now introduce bank defaults and their associated costs.

If the value V_i of a organization i falls below some threshold level $\underline{V}_i$, then i is said to *fail* and incurs failure costs $\beta_i(\mathbf{p})$. In the case of debt and equity, the default threshold is usually zero: a bank defaults when its debt obligations exceed the value of its assets (including investments, equity value, and the value of debt it is owed by others).

Once we introduce the possibility of bankruptcy and insolvency, whether a contract between ij is reflected in terms of a debt amount D_{ij} or a share amount S_{ij}will depend on bank values $\mathbf{V}$. Thus, we write $\mathbf{S}(\mathbf{V})$ and $\mathbf{D}(\mathbf{V})$. In particular, debt contracts can then be written as

$$D_{ij}(V_j) = \min \left[D_{ij}, \frac{D_{ij}}{\sum_h D_{hj}} V_j \right]$$

and equity as

$$S_{ij}(V_j) = S_{ij} \max \left[V_j, 0 \right].$$

These lead to two regimes. If $V_j \geq \sum_h D_{hj}$, organization j can repay its creditors in full such that

$$D_{ij}(V_j) = D_{ij} \quad \forall i$$

and

$$S_{ij}(V_j) = S_{ij} V_j,$$

While if $V_j < \sum_h D_{hj}$ then

$$D_{ij}(V_j) = \frac{D_{ij}}{\sum_h D_{hj}} V_j$$

and

$$S_{ij}(V_j) = 0.$$

3.4 Including Failure Costs in Market Values

The valuations in (2) have analogs when we include 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:

$$V_i = \sum_k q_{ik} p_k + \sum_{j \neq i} S(\mathbf{V})_{ij} V_j + \sum_j \left(D(\mathbf{V})_{ij} - D(\mathbf{V})_{ji} \right) - \beta_i I_{V_i < \underline{V}_i}$$

where $I_{V_i < \underline{V}_i}$ is an indicator variable taking value 1 if $V_i < \underline{V}_i$ and value 0 otherwise.

This leads to a new version of (2):

$$\mathbf{V} = (\mathbf{I} - \mathbf{S}(\mathbf{V}))^{-1} \left([\mathbf{Q}\mathbf{p} + \mathbf{D}(\mathbf{V})\mathbf{1} - [\mathbf{1}'\mathbf{D}(\mathbf{V})]'] - \mathbf{b}(\mathbf{V}, \mathbf{p}) \right), \tag{3}$$

where $b_i(\mathbf{V}, \mathbf{p}) = \beta_i(\mathbf{p}) I_{V_i < \underline{V}_i}$.[27]

3.5 Equilibrium Existence and Multiplicity

A solution for organization values in equation (3) is an *equilibrium* set of values, and encapsulates the network of cross-holdings in a clean and powerful form, building on the dependency matrix $\mathbf{A}$, as well as the debt holdings.

There always exists a solution and there can exist multiple solutions to the valuation equation (multiple vectors $\mathbf{V}$ satisfying (3)) in the presence of the discontinuities. In fact, the set of solutions forms a complete lattice.[28] As highlighted by Elliott, Golub and Jackson (2014), there are two sources of equilibrium 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 bank i is solvent and 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 set of equilibrium values forms a complete lattice, there exists a best-case equilibrium in which the set of defaulting organizations is minimal (Elliott, Golub and Jackson (2014)). The following algorithm finds the best-case equilibrium. Assuming that every bank is solvent, compute the right-hand side of (3). If any value is negative, the associated bank defaults and the values are computed again accordingly. Iterating this process yields the best-case equilibrium. There also exists a worst-case equilibrium that identifies the maxi-

[27] The number $b_i(\mathbf{V}, \mathbf{p})$ reflects realized failure costs, and is zero when failure does not occur. It always depends on the asset values through the indicator $I_{V_i \leq \underline{V}}$, but the bankruptcy costs β_i may depend on underlying asset values, $\mathbf{p}$.

[28] This can be seen by an of Tarski's fixed point theorem: by assumption banks' values depend positively on each other, which entails that failures are complements. It suffices then to check that the right-hand-side of equation (3) for bank i is weakly increasing in its own value V_i, which is true as soon when the assets of a given company are weakly increasing in all other values.

mal set of organizations that could default. It can be found using the same algorithm, but initially assuming that all organizations default.

In this paper we account for the multiplicity of equilibria as much as possible, since some results depend on which equilibrium is considered and expected by banks. In particular studying the worst-case equilibrium 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.

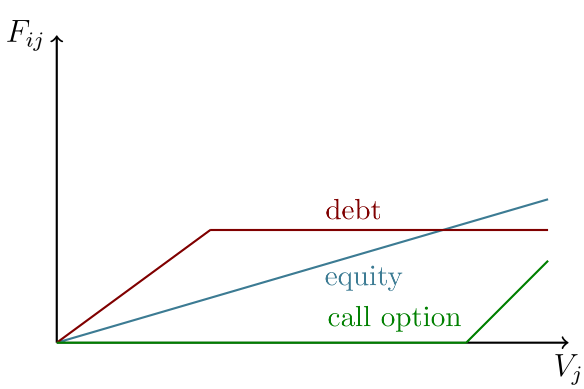
3.6 Values with more General Contracting

Consider now a financial network in which contracts are not restricted to equity and debt holdings. The book value V_i of an organization i can then be written as

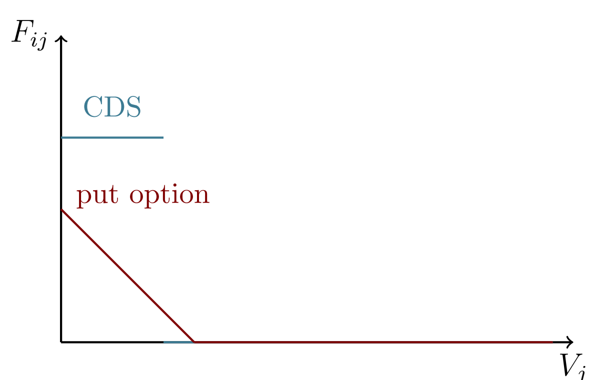
$$V_i = \sum_k q_{ik} p_k + \sum_j f_{ij}(\mathbf{V}, \mathbf{p}) - \left[\sum_j f_{ji}(\mathbf{V}, \mathbf{p}) - S_{ji}(\mathbf{V}) \right] - b_i(\mathbf{V}, \mathbf{p}), \tag{4}$$

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

A solution for organization values in equation (4) is an *equilibrium* set of values. Given that the $b_i(\mathbf{V}, \mathbf{p})$s are decreasing in $\mathbf{V}$ (higher values lead to weakly less bankruptcies) and bounded functions (supposing that the costs cannot exceed some total level), then if we take each $f_{ij}(\mathbf{V}, \mathbf{p})$ to be an increasing function of $\mathbf{V}$ (using the usual Euclidean partial order), then there exists a fixed point by Tarski's fixed point theorem for each $\mathbf{p}$. In fact, they comprise a complete lattice. Discontinuities, which come from bankruptcy costs and potentially the financial contracts themselves, can lead to multiple solutions for organizations' values. Note that assuming non-decreasing financial contracts rules out securities such as credit default swaps and put options.



(a) Non-Decreasing Financial Contracts



(b) Non-Increasing Financial Contracts

When financial contracts are not increasing functions of the values of organizations $\mathbf{V}$, there may not exist an equilibrium. As soon as some 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: 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 organizations are linked to each other through the following financial contracts: organization 2 holds debt from 1 with face value $R_{21} = 0.85$; 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 write

$$\begin{aligned} f_{21}(\mathbf{V}) &= D_{21}\mathbb{1}_{V_1 \geq \underline{V}_1} + V_1 \mathbb{1}_{V_1 < \underline{V}_1} \\ f_{23}(\mathbf{V}) &= D_{21}\mathbb{1}_{V_1 < \underline{V}_1} \\ f_{32}(\mathbf{V}) &= r\mathbb{1}_{V_1 \geq \underline{V}_1} \\ f_{13}(\mathbf{V}) &= -0.5V_3 \end{aligned}$$

with $\underline{V}_1 = D_{21}$, $\underline{V}_2 = r$, and $\underline{V}_3 = D_{21}$. First note that organization 2 and 3 never default: they have a book value of at least 2 in all cases. We then check that there is no equilibrium in which organization 1 is solvent. In such case $V_3 = 2 + r$ and $V_1 = 2 - 0.5V_3 = 0.8 < \underline{V}_1$: bank 1 defaults, which is a contradiction. Finally assume 1 defaults. Then $V_3 = 2$ and $V_1 = 2 - 0.5V_3 - \beta = 0.9 > \underline{V}_1$.

Note that in the particular case of debt and equity, financial contracts are continuous and increasing functions of the the counterparty's value.

4 Externalities and Inefficiencies in Portfolio Choices

The previous section focused on characterizing equilibrium 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.

4.1 Overly Risky Investment: The Intensive Margin

Because of interdependencies between organizations in the financial network, a bank's investment decision not only affects its values but that of others. This relates to the standard agency problem between the manager of a firm and its shareholders highlighted in Jensen and Meckling (1976). The fact that shareholders have a claim of some fraction of the firm's value changes the incentives faced by the owner-manager.[29] The later may not choose the value-maximizing actions for the shareholders. In this section, we study how financial contracts between organizations incentivize them to make overly risky investments, because they do not bear and internalize the associated costs.

Let us begin by examining the incentive problems of a single bank taking 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.

Suppose its outstanding debt D_i^L is low enough such that if the bank were to only invest in the safe asset, it would be able to pay it back: $D_i^L \leq (1+r)$. Let $s < 1/2$ be the equity share of the bank owned by outside investors. Since the residual owners hold a share $1-s$ of the bank, they form a majority and thus have control over the investment decision. Insolvency happens if the value of the portfolio falls below the bank's liability D_i^L, in which case a cost of $b \leq d$ is incurred. Under limited liability, the bank owners get a payoff of zero in case of insolvency, and the bankruptcy cost b is paid by a third party—e.g. whomever they are owed to, or the agent that steps in.

Suppose bank owners are risk neutral and choose the portfolio to maximize 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. Residual owners then solve

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

Here we do not allow for short sales.

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 q_i. This examines 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 p has a bounded and atomless distribution and 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's owners will invest entirely in the risky portfolio.*

Proposition 1 does not extend to examples in which some j is more sensitive to i's portfolio

[29]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 Hong Teoh (2009).

than i is. For instance, suppose that i is owed a large debt by j and j has equity in i and is only solvent if i's return is above $1+r$, and i has no debt and so 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, but there must be a nontrivial chance of driving a counterparty in whom i has a large stake into bankruptcy *without* having i become bankrupt.

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 costs when deciding its investment: the above maximization problem is independent of b. Indeed, under limited liability, residual owners 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.

An Example with One Bank To develop intuition further we consider the simple case of proprietary binary assets. Each bank has one unit of capital that it can either invest in its risky asset or in a safe bond. 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 $R\theta > 1+r$, 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 the risk premium $\theta R > (1+r)$, it follows directly that the bank's optimization problem has corner solution 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. 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 $\theta R - 1 + r$, 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 $\bar{q}$ or when

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

This happens when

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

Thus, there is inefficiency when:

$$0 < \theta R - (1+r) < (1+r)(1-\theta)\frac{b}{d}.$$

The bank over-invests in the risky asset when the risk premium is not high enough to compensate losses due to default. Intuitively, this is more likely to happen for higher bankruptcy costs b since the bank overlooks these costs when choosing its portfolio. The inefficient region is decreasing in R and θ, while increasing in r: the social planner is more likely to find it optimal to only invest in the risky asset for higher values of the risk premium. More surprisingly, the inefficient region is decreasing in the level of outstanding debt d and vanishes for $d \geq \frac{(1+r)(1-\theta)b}{\theta R-(1+r)}$. 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. The social planner is more likely to accept the possibility of default, and invest fully in the risky asset.

An Example with Many Banks We now give an example of investment decisions on a network of banks. Suppose the network is solely composed of debt contracts, and that each bank is *balanced* in the sense that it has as much debt coming is as coming out. Consider a bank with random incoming debt payments $\tilde{d}$, which have face value equal to the banks debt liabilities of d. The incoming debt payments may be correlated with the bank's own outcome. Assume that $D_i^A \leq R$ for all i, such that a high realization of the risky asset is enough for each bank to cover its liabilities. The bank owners choose q to maximize

$$(1-s)\left[\theta(qR+(1-q)(1+r)-d+\mathbb{E}[\tilde{d}|\theta])++(1-\theta)((1-q)(1+r)-d+\mathbb{E}[\tilde{d}|1-\theta])^{+}\right].$$

Again, regardless of the realization of $\tilde{d}$ or how it depends on the banks own payments, the optimal investment is for the bank to choose $q=1$. So, the unique investment equilibrium is for all banks to invest fully in the risky asset. Consider the case of perfectly correlated returns across banks. Then when the risky asset pays nothing the system incurs n bankruptcy costs. This leads to the same conditions for inefficiency as in the single-bank case – there is inefficiency when:

$$0<\theta R-(1+r)<(1+r)(1-\theta)\frac{b}{d}.$$

This is independent of the network structure, as long as debt is balanced debt and assets are perfectly correlated.

Next consider the case of identically and independently distributed assets. Now the probability and the extent of a default cascade depends on the network structure. For instance if the financial network forms a directed ring, then there is a default of and only if none of the risky assets pays off. Indeed, as long as one bank gets the high return, it can pay back its debt which starts a "repayment cascade". Hence the probability that all banks go bankrupt is small $(1-\theta)^n$, and otherwise the system clears. The range of parameter values for which investing fully in the risky asset is inefficient is then smaller than under full correlation of assets in that case.

4.2 Investment Incentives Under Debt vs. Equity

The result in Proposition 1 generally extends to all continuous financial contracts, which then give equivalent incentives to banks in terms of investment decisions: they make it always payoff maximizing for banks to invest fully in the asset with highest expected returns. However if a debt contract allows for discontinuous feedback effect through the value of counterparties, then it may incentivize banks to take less risk than an equity contract if they anticipate this indirect effect. The following example illustrates this intuition, and highlights the fact that incentives depend on the default-equilibrium we select—e.g. vary whether banks expect the worst-case or the best-case equilibrium to arise for a given realization of asset returns.

An Example Consider two banks $i=1,2$. Bank 1 owes $\underline{d}$ to bank 2, while 2 owes $\bar{d}$ to 1 with $\underline{d}<\bar{d}$. 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-case equilibrium—i.e. the equilibrium of bank values such that there are 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 equilibrium to arise, then her optimal portfolio is to only invest in the risky asset and set $q^* = 1$.

However in the worst-case 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

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

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

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.

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 indirect effects of one's return on its own value through the network. Despite the fact that the Modigliani-Miller theorem does not hold here, it is not clear which financial contracts are best in terms of risk-taking incentives.

What is unambiguous is that banks have lower default thresholds under equity. Keeping incentives constant, equity is then more efficient since it reduces the probability of default, and hence the expected bankruptcy costs b that the system could have to pay. 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.

The contrast between debt and equity also suggests that the systemic risk roles of institutions 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.

4.3 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 raised was whether there really was any issue of potential contagion and "dominoes", or whether the common timing of insolvencies was simply due to all banks "boiling in the same hot oil" - i.e. all having extensive exposure to an under-performing mortage market. The answer, of course, is that both were true. Banks had very correlated portfolios and all had dangerously low values in their investments, and hence most were either barely solvent, or even insolvent at the same time. Nonetheless, they also had large exposures to each other's debts, as well as to derivatives from AIG on which the counterparty could not even manage margin payments, as well as securities issued by Fannie Mae and Freddie Mac, that were both insolvent. This made it clear that a large cascade would occur without intervention.[30]

This highlights the fact correlation of investments across banks matters for financial contagion: the fact that many organization held directly or indirectly similar subprime mortgages made the whole system more fragile. 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 directly. 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 probably do not face high enough asset returns to absorb the shock.

The point that banks *prefer* to be partnered with other institutions 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 strong and stark form in our model, and so we illustrate it here. Here the correlation occurs since financial institutions prefer to choose portfolios similar to those to which they have exposure - the flip side of the choice of to whom to lend.

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. Let $s(R_i - d) > d$ such that a bank remains solvent even if its own portfolio pays zero as soon as that of its counterparty pays off. Each bank can choose in which (mix) of these two assets it wants to invest. Suppose, for now, that bank 1 is fully invested in asset 1. Then bank 2 wants to choose the same portfolio if

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

This simplifies to

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

Note that this is always true when $R_1 = R_2$, and holds even if $R_2 > R_1$ for large enough $(1+s)\theta(1-\theta)d$. 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.

If we look at the social value of the investments, no costs of bankruptcy are born if the banks are both solvent, and so having the banks choose independent portfolios rather

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

than highly correlated ones, is generally preferable. Indeed, in 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 of the assets and the other bank to invest in the other (or to have the banks hold both R_1 and R_2). Instead the banks prefer to hold all of the same investment.

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 $p_i = R$ with probability θ and nothing 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 $p_i = R$ and in which they get $p_i = 0$, subject to having a total probability of θ of getting $p_i = R$. We model this by introducing a set Ω of K equally likely primitive payoff-irrelevant states, and a set of K Arrow-Debreu securities, each paying R in exactly one state. Each bank can then flexibly choose its portfolio's state-contingent return as long as it maintains an expected return equal to θR; i.e., each bank chooses a fraction of θ of the K states that it wants to get R in. 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.[31]

Thus, if banks want perfectly correlated portfolios, they will all choose to get R in the same states, while to have independent portfolios they will all choose to get their respective R'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, 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 0's and R's that are realized. Let $p_{-i} = \mathbf{R}$ denote that all banks other than i have received R's, and $p_{-i} = \mathbf{0}$ denote that all other banks have gotten 0's.

Proposition **2.** *Suppose that all banks are solvent if they all get returns of R, 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., $V_i(p_i = R, p_{-i} = \mathbf{0}) \leq 0$ for all i), and that there is at least one bank i such that $V_i(p_i = R, p_{-i} = \mathbf{R}) > V_i(p_i = 0, p_{-i} = \mathbf{R})$. 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. Moreover, if any bank i that gets $p_i = 0$ becomes insolvent regardless of the returns of the other banks, and bank values are interconnected so that $V_i(p_i = R, p_{-i} = \mathbf{R}) > V_i(p_i = R, p_{-i} = \mathbf{p})$ for each i when $\mathbf{R} \neq \mathbf{p}$, then there is a* unique equilibrium *out of all possible portfolio configurations and it involves perfect correlation.*

[31] 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 in, so they cannot, for instance, choose to get $2R$ in some states. As will become clear in the proof, the ideas generalize.

Note that the condition $V_i(p_i = R, p_{-i} = \mathbf{R}) > V_i(p_i = 0, p_{-i} = \mathbf{R})$ is extremely weak, and necessary to get any results. If it was true that all banks see no increase in value when their portfolio changes, even in the best possible state in terms of the payments made by other banks in the network, then banks will not care about their returns at all. Everything would be an equilibrium. That would require that essentially all returns nR in the full network are owed as debt to outsiders, and so no bank in the network ever has any equity value. The condition that $V_i(p_i = R, p_{-i} = \mathbf{0}) \leq 0$ for all i is also weak in that it examines the extreme case where all other institutions in the whole economy get no return, and requires that it cannot survive. This condition is not necessary for the result (e.g., see the example above), but makes the argument transparent.

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 gain from uncorrelated portfolios.

The condition that $V_i(p_i = R, p_{-i} = \mathbf{R}) > V_i(p_i = R, p_{-i} = \mathbf{p})$ for each i when $\mathbf{R} \neq \mathbf{p}$ does not require that all banks be directly linked to each other in terms of debt or equity, just that there be a directed path of equity and/or debt defaults between any pair of banks that lowers the one's value when the other defaults.

[[fill in sufficient conditions for the above]]

[[Fill in social optimum and discussion]]

4.4 Too Few Partners: The Extensive Margin

As mentioned above, 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. However, it also leads bank portfolios to become more diversified through cross-holdings, reducing the risk of a first-failure. Beyond a certain level of financial integration, this second effect dominates. There is then a critical number of counterparties at which systemic risk is maximal, what Elliott, Golub and Jackson (2014) call the sweetspot.

In the previous sections, we looked at the intensive margin of investment choices, and highlighted the over-investment in risky portfolio. We now turn to the extensive margin and study how many counterparties a bank optimally chooses. We here get under-diversification of bank portfolios.

In particular, consider a set of n of similar banks, each with its own investment portfolio p_i, and let us suppose that they each have an equal expected value and they are exchangeable random variables. Banks can enter into swap arrangements – effectively granting each other equity in their portfolios, and equalizing their holdings. For instance, two banks in net would end up with book values of $\frac{1}{2}p_i + \frac{1}{2}p_j$ (less other holdings) each, instead of p_i and p_j. This can change the probability that they remain solvent and so increase their expected payoffs.

Let's consider a bank's choice between two different regimes: one in which it swaps with m other banks and another where it swaps with $m-1$ other banks. Without loss of generality, we write the problem from bank 1's perspective and consider partnering with the first m banks. Each bank has a debt $D_i^L = d$ and no other contracts. It costs c to contract with another banks.

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

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

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 to $m-1$ if and only if:

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

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

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

If the p_i's are less then perfectly correlated, then

$$\Pr\left[\sum_{i=1}^{m-1}\frac{p_i}{m-1}\geq d\right]-\Pr\left[\sum_{i=1}^{m}\frac{p_i}{m}\geq d\right]<0$$

and so the planner always has a weaker inequality, and prefers higher values of m. For high levels of bankruptcy costs b, this gap becomes larger.

The effect of asset correlation on the optimal number of partners is ambiguous, as there are two competing effects. First, as asset correlation increases, swapping with an extra bank does not diversify one's portfolio that much and the difference in the probabilities of default between m and $m-1$ partners decreases. However the effect on the marginal change in portfolio return conditional on solvency may be of ambiguous sign. In any case, as assets become more correlated, the decision rule of individual banks characterized by the first inequality becomes closer to that of the planner: indeed the gains from diversification are lesser from the planner since the reduction in the probability of default from diversifying bank portfolios is smaller.

Thus, as soon as assets are not all perfectly correlated, banks under-diversify their portfolio in terms of the number of their counterparty relationships in addition to choosing overly risky investments.

4.5 Inefficient Bank Size

Can mergers reduce inefficiencies in the financial network, either by internalizing some of the externalities banks have on others or by making the network more resilient to risk?[32] Let N be a set of banks, $\mathbf{D}$ the matrix of debt claims, and $\mathbf{S}$ the matrix of equity claims. We look at the effect of a merger between bank i and j into a larger organization k such that $D_{kl} = D_{il} + D_{jl}$ and $D_{lk} = D_{li} + D_{lj}$ for all $l \neq i, j$, and similarly for equity shares.

How bank size interacts with the fact that banks choose overly risky investments (highlighted in section 4.2) 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 4.3.1. with $n = 3$ banks. Bank 3 can only invest in the safe asset; it has equity share s in bank 1 and debt claim $\underline{\mathrm{d}}$ in bank 2. It owes $\bar{d} > \underline{\mathrm{d}}$ to bank 2 as well. Suppose $1 + r + \underline{\mathrm{d}} < \bar{d}$ 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 $\varepsilon := \bar{d} - \underline{\mathrm{d}} - 1 - r$ 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

[32] See Kanik (2018) for a detailed discussion of banks' incentives to merge to save themselves from failure in a cascade. Here we examine banks' incentives to merge well before a cascade.

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.

5 Evaluating and Addressing Systemic Risk: Flying Jets without Instruments

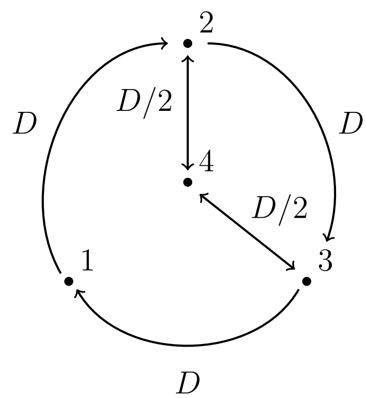
As should be obvious by now, properly assessing systemic risk involves a holistic view of the network. Contagion can occur directly between counterparties, or even indirectly without intermediate defaults, as banks' values are interdependent even at a distance via chains of relationships. We first illustrate this point with an example, and then discuss issues of ensuring solvency.

Attempting to assess systemic risk without 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.

5.1 The Necessity of Network Information

In particular, to assess systemic risk regulatory authorities use stress tests, which are either run by the authority itself or by individual banks, usually 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. Balance sheet data does not usually provide any information on the identity of each counterparty and its other investments and counterparties, and hence on the network structure. We here argue that any measure of systemic risk solely based on such "aggregate" data can be misleading and 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 an obvious one, but worth emphasizing given its importance.

For simplicity, consider a network in which banks only have debt contracts between each other. If a "good" measure of systemic risk based on aggregate data existed, then in this case it would only depend 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 balance sheet data solely. Consider the balanced network composed of four banks depicted in the following figure:



Suppose the portfolio of bank 1 and 4 yields $-\varepsilon$ such that they are both insolvent, whereas all the other organizations earn returns on their investment in $(D/2, D)$. Note that bank 1 and 4 have the same balance sheet since $D_1^A = D_2^A = D$ and $D_1^L = D_2^L = D$. However only bank 1 induces widespread default contagion if it remains insolvent. Indeed, bank 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 to go bankrupt, while bailing out bank 4 does not change anything: balance sheet data is not enough to distinguish two financial organizations like these.

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 about half 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 a nontrivial network of systemic risk are beginning to emerge (e.g., see Covi, Gorpe and Kok (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.

5.2 Financial Centrality

We now provide a measure of financial impact of a given organization that relies on network information.

Let us define the impact that some institution has on the rest of the economy by a change in its portfolio. Within our model, there is an obvious – and thus essentially unique – way to define a firm's impact on the rest of the economy. 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 to be

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

This is the total impact on the economy to the ultimate shareholders and debt-holders that comes from a change in i's portfolio on the rest of the economy.[33] Note that this is equivalent to

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

We can define the *net financial centrality* of i is

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

which is the impact beyond the direct drop in i's portfolio value, and 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 default centrality of i is 0.

Implicit in defining financial centrality, one has to take a stand on which equilibria are being examined since those define the values $V(\mathbf{p})$ and the bankruptcy costs $b(\mathbf{p})$. Typically we are interested in either the best-case or worst-case 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).

In a balanced network in which assets exactly compensate liabilities on every bank's balance sheet, when considering the best equilibrium no bank has any net financial centrality. 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 could still be very fragile. For instance, consider a balanced network with no contracts other than debt. One can change many 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 default centrality of a bank could be large. More generally, this implies that (net) financial centralities are discontinuous and can be very sensitive.

In a world with only debt between banks, this measure will follow chains of cascading defaults from the original bank. Some banks can stop such chains if they have enough value and/or small enough debt out compared to in— $D_i^A \geq D_i^L$;—these are *buffer* banks. Note that these "chains" could hit some banks multiple times and so intersect. With equity, these chains 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 defaults. This seems well-defined and in principle easy to calculate if one knows the network, as the best-case equilibrium calculation is a linear time algorithm (at most n steps).

[33]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.

5.3 Bailouts

5.3.1 Balanced Debt

First, it is useful to examine a special case, that we refer to as *balanced-debt network*.

We say that there is *balanced debt* in the system if $D_i^A \geq D_i^L$ for all i. This is a requirement that each bank has at least as much debt coming in as going out. Since all debt claims and liabilities must cancel out on aggregate, this implies $D_i^L = D_i^A$ for all i. Under this requirement, as long as a bank receives the debt payments it has coming in, then regardless of its investments, it will be able to meet its debt obligations going out. Hence there always exists a best equilibrium in which no-one ever defaults, since if all banks but i honor their debt contracts then i can also pay back its debt fully in a balanced network.

This logic holds irrespective of the network structure, and even when all risky assets have a zero realization. However in that case, there exists another equilibrium in which everyone defaults and the system incurs bankruptcy costs for all organizations that have any debt liabilities. One can interpret this worst equilibrium as a failure to coordinate 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, as it would require all write-offs to be done simultaneously to maintain solvency. Moreover, in practice the debts will 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, but we will discuss both equilibria in what follows.

The possibility of such repayment cascade in a balanced network can be exploited by a regulator that tries to minimize default costs. Suppose the regulator can inject liquidity into the network, to allow some banks to remain solvent—i.e. it can bailout a subset of banks $B \subseteq N$. If bank $i \in B$ is bailed-out, it pays back its debt to all its creditors, who then may become solvent themselves. Depending on the network structure, bailing-out a few banks may make the entire system solvent even when a crisis occurs and none of the risky asset pays off. 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.

A Simple Example. Consider a balanced ring network composed of n banks, such that bank i owes debt d to bank $i+1$. To close the ring, bank n has liabilities towards bank 1. Note that even when all portfolios yield zero, there exists an equilibrium in which all banks remain solvent. Such equilibrium may be unlikely to arise spontaneously, but can be induced by bailing-out a single bank. Indeed suppose the regulator 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 unravels up to bank n who repays bank 1. The regulator recoups its initial liquidity input at the end of the cascade, and is thus able to prevent n defaults at no explicit cost.

If banks are in a more complex network, then the minimal set of banks that need to be saved in order to make the system liquid is more complicated. The optimal strategy for the regulator is to exploit cycles in the network, so as to induce as many cascades as possible.

The minimal set obviously depends on the repayment rule in case of default. Consider the strong requirement that a bank can pay back its debt if and only if either of the two following conditions hold: the bank is bailed-out or *all* of its incoming debts are first paid back. In particular this rules out partial payments: even if a bank has some money coming in, she cannot use it 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 is always larger under this rule than when partial repayments are allowed.

PROPOSITION **3.** *Consider a balanced network. Then banks on the network end up being solvent in the worst equilibrium if and only if there is at least one saved or unilaterally solvent bank on each directed (simple) cycle.*[34] *All banks are solvent in the best equilibrium.*

In the case in which there are some cycles with insolvent banks, then one can consider the problem of injecting the minimum amount of capital in order to ensure solvency of the full network in the worst equilibrium. We describe such an algorithm for the more general case where not all organizations are balanced, of which this is a special case.

Finding a minimal such set is computationally hard and requires a lot of information on the network structure, as the regulator first has to identify all cycles and then compute the minimal set that overlaps with all of them. The first step by itself is already computationally demanding in a large and diverse network (i.e. NP-hard).[35] Furthermore note that under the weaker rule that allows partial repayments, bailing-out one bank from each directed cycle is not always necessary to ensure solvency of the system.

5.3.2 Imbalanced Portfolios

We now allow bank portfolios to be unbalanced, which in particular means a bank may have more debt liabilities than assets. Let c be a directed simple cycle of g, and denote by $N(c)$ the set of nodes that belong to c. Finally let $s_c(i)$ and $p_c(i)$ be respectively bank i's successor and predecessor in cycle c. Then a simple cycle is a list of links $p_c(i)i$, $is_c(i)$ such that $D_{i,p_c(i)} > 0$, $D_{s_c(i),i} > 0$ for all $i \in N(c)$, and such that there is no strict subsequence that also forms such a directed cycle.

To understand what happens at a given value of portfolios $\mathbf{p}$, let us define $\widehat{V}(\mathbf{p})$ to be the values of all institutions when we presume that all debts are paid and no bankruptcy

[34]A simple cycle is one that only contain 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.

[35]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).

costs are incurred. This is not an equilibrium calculation, since it can still be that some of these values are negative, and so without further intervention, then some of the banks would default. In order to have a fully solvent system, any negative value has to be brought up to be positive. However, some of these are inter-dependent, so the lowest payment to ensure full solvency is less than $\sum_{i:\widehat{V}(\mathbf{p})<0} -\widehat{V}(\mathbf{p})$. The problem for figuring out the minimum capital that needs to be injected to ensure full solvency is

$$\min_{\mathbf{p}':\widehat{V}(\mathbf{p}')\geq 0} \sum_i |p'_i - p_i|.$$

The following proposition shows how this works in a network with only debts and no equity. In that case, let the *net imbalance* of bank i be $[D_i^L - D_i^A - p_i]^+$. This is the total excess debt that the bank owes compared to what it is owed. In order to ensure that each bank is solvent in the best equilibrium, it becomes clear that the necessary total payment is the sum of these payments. To ensure solvency in the worst equilibrium, we also need that there is at least one bank on each cycle for which they can make their payments unilaterally, so that $p_i \geq D_i^L$. The minimum payment is thus the total net imbalance, plus whatever additional minimum payment is needed to get at least one unilaterally solvent bank on each cycle.

Proposition **4.** *Consider a imbalanced network in with no equity. Then banks on the network end up all being solvent in the best equilibrium if and only if each bank has a net imbalance that is 0, and the necessary payment needed if some bank is insolvent is the total net imbalance in the economy. There is solvency in the worst equilibrium, if and only if the total net imbalance is 0, and there there is at least one saved or unilaterally solvent bank on each directed simple cycle.*

To ensure solvency of the entire network, the regulator first has to rebalance the portfolio of all banks that are net borrowers. Then we are back to the balanced case, and a bank in each directed simple cycle must be solvent unilaterally if we have to deal with the worst equilibrium. It is hence necessary and sufficient for the regulator to inject the net imbalance of all banks *plus* ensuring a unliaterally solvent bank on each cycle.

It can hence be significantly costlier to ensure solvency of an imbalanced financial network since the regulator necessarily has to inject the amount of net debt of all net borrowers. In reality, most institutions 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 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.

Given the potential complexity and cost of maintaining the whole system solvent in a crisis, what requirements should be put on bank portfolios to prevent default cascades? How much should they hold in a risk free asset, or equivalently have in reserve? Traditionally, banks have been regulated to hold enough assets in reserve to be able to meet their short-term debt or demand deposit obligations.[36] Those deposit requirements are far short of all of the debts that are held, especially once one moves beyond traditional banks. If banks face low or negative returns to investment, these requirements are thus not enough to prevent a first failure, and a potential contagion. To reduce the cost of bailouts, the regulator could set requirements high enough to make the network balanced, but low enough to allow fruitful investments. This would guarantee that each organization would have enough reserves to pay its *net* debt, but would not prevent self-fulfilling runs and default cascades.

6 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 very difficult politically to get. 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.

The second point is that the externalities in financial networks lead to several incentive problems: institutions have incentives to take overly risky positions, to involve too few counterparties, and to overly correlate their portfolios with those of others. These are harder to fix. Excessive risk is partly, but imperfectly, addressed by reserve requirements. The imperfection relates back to the fact that the systemic risk is not appropriately measured, and crude reserve requirements will generally either fall short or be excessive, and will only occasionally be at appropriate levels. Moreover, such reserves are only imposed based on a portion of the liabilities and only for a subset of financial institutions. Incentives to take on too few counterparties and to overly correlate portfolios are also issues that have generally been completely ignored, and not ones for which there are easy policies. One could require markets to have Central Counterparty Clearing Houses - CCPs. These pass all transactions through a central intermediary, which then can monitor positions and impose margin requirements. One then has to worry about providing the CCPs with appropriate

[36] In the US, depository institutions are required to hold between 0% and 10% of their deposit liabilities in reserves, the precise ratio depending on the amount of its liabilities.

incentives and worry about their size.[37] 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. 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.

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, 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.

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.

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.

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

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.

Bruche, Max and Carlos Gonzalez-Aguado, "Recovery rates, default probabilities, and the credit cycle," *Journal of Banking & Finance*, 2010, *34* (4), 754–764.

Brunnermeier, Markus K, "Deciphering the liquidity and credit crunch 2007-2008," *Journal of Economic perspectives*, 2009, *23* (1), 77–100.

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

Caballero, Ricardo J and Alp Simsek, "Fire sales in a model of complexity," *The Journal of Finance*, 2013, *68* (6), 2549–2587.

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

Centola, Damon, *How Behavior Spreads: The Science of Complex Contagions*, Vol. 3, Princeton University Press, 2018.

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.

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.

Diamond, Douglas W and Raghuram G Rajan, "Fear of fire sales, illiquidity seeking, and credit freezes," *The Quarterly Journal of Economics*, 2011, *126* (2), 557–591.

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

_ , Andreas Eckner, Guillaume Horel, and Leandro Saita, "Frailty correlated default," *The Journal of Finance*, 2009, *64* (5), 2089–2123.

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*, 2018.

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

Fan, Yuan, Gang Feng, and Yong Wang, "Weight balance for directed networks: Conditions and algorithms," in "Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on" IEEE 2010, pp. 269–274.

Farboodi, Maryam, "Intermediation and voluntary exposure to counterparty risk," *mimeo:*, 2014.

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.

Fleming, Michael J and Asani Sarkar, "The failure resolution of Lehman Brothers," *Economic Policy Review, Federal Reserve Bank of New York*, 2014, *20* (2).

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.

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., *Social and economic networks*, Princeton: Princeton University Press, 2008.

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

_ **and Asher Wolinsky**, "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 1996, *71* (1), 44–74.

— **and Evan C. Storms**, "Behavioral Communities and the Atomic Structure of Networks," *Arxiv: https://arxiv.org/abs/1710.04656*, 2017.

— **and Stephen Nei**, "Networks of Military Alliances, Wars, and International Trade," *Proceedings of the National Academy of Sciences*, 2015, *112(50)*, 15277–15284.

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.

Keynes, John Maynard, *The General Theory of Employment, Interest, and Money*, London: Macmillan, 1936.

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.

Morris, Stephen and Hyun Song Shin, "Unique equilibrium in a model of self-fulfilling currency attacks," *American Economic Review*, 1998, pp. 587–597.

Reinhart, Carmen and Kenneth Rogoff, *This time is Different*, Princeton University Press: Princeton, 2009.

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

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.

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.

7 Appendix

Proof of Proposition 1: Let μ be the measure on p, the vector of all portfolio values, and let

$$A(q_i) = \{p \mid q_i p_i + (1 - q_i)(1 + r) + \sum_{j \neq i} s_{ij} V_j(p, q_i) + d_i^{in}(p, q_i) > d_i^{out}\}.$$

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

Consider any $q < 1$ for which $\mu(A(q)) > 0$ (which are the only possible optimizers), and let us examine the gain in utility that results from increasing q to $q + \varepsilon$. We show that for any such q 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

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

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

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

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

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

Thus,

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

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

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

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

which is at least

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

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

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 requires the existence of at least one state of the world in which $p_i = R$ but $p_j = 0$ for all $j \neq i$, and similarly of at least one state in which $p_i = 0$ but $p_j = R$ 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

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

First note that

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

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

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

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 in 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 $V_i(p_i = R, p_{-i} = 0) = 0$ for all i. Again, the incentive condition boils down to

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

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

Finally, suppose that any bank getting a return of 0 becomes insolvent, and that every bank satisfies $V_i(p_i = R, p_{-i} = \mathbf{R}) > \mathbf{V_i}(\mathbf{p_i} = \mathbf{0}, \mathbf{p_{-i}} = \mathbf{R})$. It is sufficient to show that there are no other equilibria (since we have just established that perfectly correlated portfolios is an equilibrium, even without requiring the additional assumptions). Consider some configuration that is not perfectly correlated. Then there must exist some bank i getting $p_i = 0$ when $p_{-i} = \mathbf{p} \neq \mathbf{0}$. We show that

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

By assumption $V_i(p_i = 0, p_{-i} = \mathbf{p})$ and $V_i(p_i = 0, p_{-i} = \mathbf{R})$ are both no more than 0, and so it is sufficient that

$$V_i(p_i = R, p_{-i} = \mathbf{R}) > \mathbf{V_i}(\mathbf{p_i} = \mathbf{R}, \mathbf{p_{-i}} = \mathbf{p}).$$

▮

Proof of Proposition 3: 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 a balanced network can be decomposed into a set of strongly connected components (in which each bank can reach each other one by a directed path), such that there are no directed paths pointing in or out of any of these components (e.g., see Fan et al. (2010)).

Thus, it is without loss of generality we prove the claim for a strongly connected component.

Second, we prove that if there is some directed cycle that has no unilaterally solvent bank on it, then none of the banks end up being solvent. [[insert proof: even if all other banks outside of this cycle are solvent, none of these banks will ever be fully paid back.]]

Finally, we prove that if each directed cycle has at least one unilaterally solvent bank on it then all banks are solvent. We prove this by induction on the number of banks n.

It clearly holds for $n = 2$ since then one bank being solvent means that the other gets all of its incoming debts and can pay the full amount out.

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* saved/solvent in the network with n nodes.

Let X_1^{in} be the set of nodes that only point into 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:

(i) $X^{out} = N \setminus \{i_0\}$. In that case saving i_0 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. Since $X^{out} = N$, the system clears, and the claim holds.

(ii) $X^{out} \neq N \setminus \{i_0\}$.

Let us consider the subgraph found by removing $X^{out} \cup X^{in} \cup \{i_0\}$. Note that this must contain at least two banks, and at least one cycle, as otherwise all the debt from at least one bank would flow into X^{in}, which would be a contradiction.

Note that the subnetwork $N \setminus X^{out} \cup X^{in} \cup i_0$ is of a structure that all edges lie on directed cycles, and still has a solvent/saved bank on each simple directed cycle since the only simple directed cycles that were removed were ones that included i_0, and the ones that did not include i_0 all remain (by removing $X^{out} \cup X^{in} \cup i_0$ we removed all cycles to which i_0 belongs, but none of the other cycles). In particular, note that all remaining banks are part of cycles to which i_0 does not belong. By contradiction, supposed that one of the remaining bank j is only part of cycles that contain i_0. Then j must be point to banks that are only part of cycles containing i_0, and similarly those banks must be only pointing at similar banks, etc. This means that $j \in X^{in}$ which is a contraction. Hence the remaining subnetwork of size at most $n-1$ can be balanced, and when balanced can be saved by bailing-out a set B that has one bank per cycle by assumption.

We claim that bailing-out $B \cup i_0$ is enough to ensure solvency of the original network. First note that bailing out i_0 ensures solvency of X_1^{out} after one round, then X_2^{out}, and eventually all of X^{out}. Then recall that X^{in} only points at bank in $X^{in} \cup i_0$, so they do not owe money to the rest of the network. Now consider the subnetwork $N \setminus X^{out} \cup X^{in} \cup i_0$, in which the debt either comes from X^{out} or from within the subnetwork. All debt from X^{out} is paid back. Bailing out B then ensures than all debt is paid back in this subnetwork (noting that by definition none of the banks in X^{in} could have debts to any of the remaining banks in this subnetwork). By construction, 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. ∎

A Note on the Existence of Equilibrium Values from Equation (2): $(\mathbf{I} - \mathbf{S})$ 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. ∎