Entangled Financial Systems *

Adam Zawadowski
Princeton University
azawadow@princeton.edu

June 16, 2009

Abstract

Counterparty risk was one of the main reasons cited for government intervention in the crisis of 2008. In my model, the financial system is "entangled" and thus susceptible to counterparty risk because financial institutions hedge risks using a network of bilateral over-the-counter contracts. If market participants have few large counterparties, they choose not to insure against a low-probability counterparty default, even if insurance would be socially optimal. Thus the collapse of a single market can lead to a complete collapse of the financial system. Contagion spreads solely through the loss of hedging contracts: there are no direct credit exposures between participants. I show that even though the sparse network structure of hedging contracts is central to the externality, participants are not willing to diversify if that is costly to do. I show that inefficiency is most likely in the early stages of the development of bilateral hedging contracts. The model implies that in case of inefficiency, regulatory intervention making insurance against counterparty risk mandatory, e.g. through a well capitalized clearinghouse, is welfare improving.

*I thank Markus Brunnermeier and Wei Xiong for advice, and Amil Dasgupta, Thomas Eisenbach, Zhiguo He, Jakub Jurek, Nobuhiro Kiyotaki, Péter Kondor, Stephen Morris, Tamás Papp, José Scheinkman, Felipe Schwartzman, Hyun Shin, Ádám Szeidl, Béla Személy, Balázs Szentes, Motohiro Yogo and seminar participants at the Federal Reserve Bank of New York, Princeton University macro, micro and finance workshops for useful discussions and comments on this and earlier versions of the paper. All remaining errors are mine. First draft: April 3, 2008.
For the latest draft see: www.princeton.edu/∼azawadow

1 Introduction

Modern financial institutions are highly leveraged and entangled in a network of illiquid bilateral hedging contracts. Most of these contracts do not even show up on the institutions' balance sheets, such as over-the-counter (OTC) derivatives. The fear of these instruments effecting the whole world economy was a major argument brought up when the Fed helped save Bear Stearns in March 2008[1]:

> "Bear is a counterparty to some $10 trillion of over-the-counter swaps. With the broker's collapse, the fear that these and other contracts would no longer be honored would have infected the world's derivatives markets. Imagine those doubts raging in all the securities Bear traded and from there spreading across the financial system; then imagine what would happen to the economy in the financial nuclear winter that would follow. Bear Stearns may not have been too big to fail, but it was too entangled."

In this paper, I propose a model of the core of the financial system in which institutions get endogenously entangled in bilateral OTC contracts to hedge their exposure. I assess whether the failure of a single institution or market could indeed lead to the collapse of the whole system. The main goal is to pin down the conditions under which financial institutions have an incentive to get exposed to counterparty risk without hedging it. Understanding such systems is crucial in order to construct efficient financial regulation in the future.

In my baseline model, financial institutions (banks from hereon) specialize in different markets and have offsetting risks. These uncertainties can be hedged by bilateral contracts linking all banks together. If equity is costly and borrowing cheap, banks choose to finance their activities by short-term borrowing, keeping as little capital as possible to satisfy incentive constraints. In order to maximize leverage, banks even use their non-pledgable payoff linked to their survival, such as long-term profits or reputation, as "collateral" to overcome the moral hazard problem, even though these cannot be seized by the lender. In effect, these incentive constraints mean that the equity of the financial institution has to satisfy a Value-at-Risk constraint.

In the event that one of the markets is hit by a large negative shock, the bank specializing in that market goes bankrupt, leaving its counterparties (neighbors from hereon) unhedged. Higher risk in these neighboring banks increases their probability of default, thus decreasing expected value of non-pledgable payoff. Since this is used as "collateral", its decrease makes the bank's short-term lenders wary of its incentives, withdrawing short-term funds and leading to liquidation of these neighboring banks too. In a reduced form framework this means that banks that have lost their hedge do not have enough equity to satisfy their VaR constraint any more. If all the banks are linked, this mechanism leads to a complete collapse of the financial system through sequential withdrawal of short-term funding.

The basic question is why do banks not insure against the failure of the counterparty if that could have devastating consequences? I show that individual decisions endogenously lead to a contagious

[1] "Wall Street's crisis", The Economist, March 22, 2008

network, even though this is socially suboptimal. The intuition is as follows: it is costly for every bank to individually stabilize the financial system, so the equilibrium becomes one in which every institution wants the other one to pay the cost of stability. The externalities through bilateral contracts lead to a market failure: the public good of financial stability is not provided by the market. This inefficient contagious equilibrium exists for a realistic set of parameters, in particular for relatively rare but devastating crisis events.

The network structure of bilateral claims is crucial for the externality. If instead all risks were traded on centralized markets through which the risks would be evenly spread over the system, all banks would fully internalize the effect of the bankruptcy of a counterparty. The network structure implies that every bank has only a few large counterparties, so banks only care about the potential bankruptcy of their neighbors when making decisions. Furthermore they do not take into account the effect of their decisions on others: their neighbors, the neighbors of their neighbors, etc. Another contribution of this paper is to show that even with small costs to diversifying counterparty exposure, e.g. by setting up markets for risks, the equilibrium with large counterparty exposure is stable.

The result that banks fail to insure against counterparty risk gives an insight on why participants in the financial market have resisted the call for setting up a central counterparty for CDS contracts. This counterparty would basically act as an insurer, making the trades more expensive. The model shows that there is scope for government intervention, as recently argued, among others, by Darrell Duffie.[2] First, "regulators should press dealers to clear more types of derivatives with the same clearinghouse". Second, "regulators should ensure that a clearing counterparty is extremely well capitalized and has strong operational controls".

The model in this paper is generalized to incorporate the case of a large central counterparty. The analysis shows that regulatory intervention to insure the strong capitalization of the central counterparty is crucial since it fails to internalize all the benefits of a strong balance sheet. If it goes bankrupt, all of its counterparties fail since they do not find it optimal to hold extra capital for this low probability event.

There has been substantial debate whether saving Bear Stearns was necessary. On the other hand, many blame the financial turmoil of September 2008 on letting Lehman Brothers fail. One of the basic arguments for letting Lehman fail was that counterparties had enough time and incentives to eliminate any counterparty exposure to Lehman. My model shows that this argument cannot be applied in general, since substantial part of the benefits from hedging counterparty risk are systemic and are not internalized by the individual bank.

One way to insure against counterparty risk is to buy default swaps on the bank one has bilateral contracts with. The idea of banks not holding enough default swaps seems surprising at first in a market where, many argue, there are already way too many credit default swap contracts outstanding. In my model, not having enough default swaps refers to default swaps of financial institutions on

[2] "Derivatives and Mass Financial Destruction: Complex financial markets can be useful if regulated properly" by Darrell Duffie, The Wall Street Journal, October 22, 2008

other financial institutions, with whom they do bilateral contracting. The events around the default of Lehman Brothers yields some evidence that counterparties did not ex ante have enough insurance against the default of the banks they did business with[3]:

> "Spooked that other securities firms could fail, hedge funds rushed to buy default insurance on the firms with which they did business. [...] The moves dramatically drove up the cost of insurance on Morgan Stanley and Goldman Sachs debt"

The mechanism proposed in this paper by which contagion spreads is, to my best knowledge, also novel. Previous work on financial contagion concentrated on direct credit exposure to explain contagion.[4] However, survey evidence shows that most of the credit exposure through OTC contracts is covered by collateral.[5] In this paper, I show that even if there is no direct loss when a neighbor defaults, there is substantial loss from becoming unhedged that can lead to contagion. The aftermath of the Lehman bankruptcy shows that it is very costly to rehedge risks in case of crisis.[6] To capture this insight in the model, I assume that rehedging is impossible once the natural hedging counterparty goes bankrupt. The results, however, carry over for the case where rehedging is possible but only at a high cost.

The results also shed light on how an idiosyncratic shock in a single market can lead to a large drop in overall real output, a question that has long puzzled macroeconomists. The financial system with its high leverage and binding incentive constraints works as an amplifier. Linkages between banks in different markets make the crisis spread to otherwise unrelated markets. Markets that otherwise would not move together at once become correlated in crisis.

Another important insight is that these results hold despite the existence of ample liquidity outside the financial system. However, investors choose to buy government bonds instead of bonds issued by financial institutions. Whether the idiosyncratic shock in a single market becomes a systemic crisis is endogenous: it depends on the equity choice of market participants. The main point is that outside liquidity cannot be drawn upon by the banks because of asymmetric information and moral hazard considerations. Thus the model also gives an explanation for "flight to quality" in crisis: investors move from lending to the financial sector to lending to the government.

The extension of the model to a financial development context also shows that in the early stages of financial development, the network of bilateral claims is endogenously sparse, thus likely to be inefficient. It is exactly after the initial boom years that the financial system is most susceptible to exogenous shocks. Only in the later stages of financial development does the network become dense and thus in effect become a complete market without inefficiency.

[3] "Lehman's Demise Triggered Global Cash Crunch", The Wall Street Journal, September 29, 2008

[4] See for example Allen and Gale (2000) and Kiyotaki and Moore (1997).

[5] ISDA Margin Survey 2008, a survey conducted among the most important participants in the OTC market by the International Swaps and Derivatives Association, reports that approximately 65% of OTC exposure is covered by collateral or margin accounts and this number is even higher among the largest market participants.

[6] "The great untangling", The Economist, November 8, 2008

The paper is structured as follows. Section 1.1 discusses how the paper relates to previous work. Section 2 shows the main inefficiency result in a reduced model of the baseline model with exogenous VaR constraints. The full microfounded model is presented in detail in Section 3: it endogenizes both the VaR constraint and short-term debt. Section 4 then puts the reduced form model in a general equilibrium framework where the network structure and the return on equity is endogenous. Section 5 analyzes an asymmetric network structure with a large financial hub. Section 6 concludes and discusses avenues for further research. All proofs not included in the text are relegated to the Appendix.

1.1 Related literature

The two most closely related models are Allen and Gale (2000) and Kiyotaki and Moore (1997). Both of the papers model contagion through loans or credit which, however, seem unlikely to cause contagion in a modern financial system where most loans are collateralized, with margin accounts adjusted regularly and interbank deposits not that important any more.

Allen and Gale (2000) model contagion through interbank deposits in an unlikely event of an aggregate liquidity demand shock. The main imperfection is that interbank assets are not state-contingent, which contrasts to my model where I allow for a wide range of state-contingent contracts. A similarity is that the incomplete network structure of bank ties is crucial for contagion, however, this is simply assumed in their paper, its source is not modeled. The authors themselves admit that their model is more a model of "banking crises in the United States in the late nineteenth and early twentieth centuries". Furthermore, the probability of crisis in their setting is zero, thus no welfare analysis is possible.

In Kiyotaki and Moore (1997), bilateral links result from the specificity of intermediate goods, not from hedging of endowment risk. Thus their setup is hard to interpret in a financial context. Inefficiency arises since claims cannot be netted out due to the specificity of intermediate goods which are used as collateral. In modern financial systems most loans are collateralized using liquid assets, thus their model is more a model of trade credit in supply chains. In their setting, entrepreneurs are credit-constrained, while in my model low equity is a choice of the banks. The authors show that under some circumstances agents chose not to insure against bad states of nature, since they are unwilling to give up costly equity today. However, they do not compare the privately optimal decision to the socially optimal one.

An insight related to this paper is that of Brunnermeier (2009), who argues that in case of offsetting OTC contracts, worries about a single counterparty can lead to systemic crisis. He does not formalize his argument, or propose a model, but his argument is different in several aspects. In that setting, knowing the whole network would help overcome a systemic crises, while in my model this is not the case. This is due to the fact that in my model contracts are not perfectly offsetting and participants specialize in different risk markets. This means that a netting arrangement such as proposed by the author cannot eliminate the risk of contagion in my model. A similar inefficiency result is derived in

Caballero and Simsek (2009). There, uncertainty about the network structure and high risk-aversion lead to banks over-hoarding liquidity in case they know one bank has been effected.

Other papers argue for establishing a centralized counterparty (CCP from hereon) for OTC derivatives too, however for different reasons than that brought forth in this paper. Duffie and Zhu (2009) shows that if all OTC derivatives were traded through the same clearinghouse, that would economize on costly margin accounts. In Acharya and Bisin (2009) a central clearinghouse would be welfare improving since it would limit unobserved excessive positions in otherwise opaque OTC markets. On the other hand, Pirrong (2009) lists several reasons for a potential negative effect of introducing a centralized clearinghouse: a CCP has less incentive to monitor counterparties and dealers might take advantage of potential mispricings to

There are other papers that model externalities in credit markets. Lorenzoni (2008) shows that in a competitive setting there is inefficient overborrowing, since agents do not internalize their effect on others in the market through fire-sale prices in case of forced liquidation following a negative productivity shock. While this model does show externalities exist in crisis, it cannot explain how small idiosyncratic shocks lead to macroeconomic fluctuations. This is due to the fact that given centralized markets, risks are spread over the system, thus one does not get a cascade of defaults that amplify initial shocks.

2 Reduced form model

I first introduce a reduced form model to show the main source of inefficiency in an entangled financial system. Section 3 then presents a microfounded model that rationalizes the assumptions in this reduced form version of the model. In the generalizations of the model, in Sections 4 and 5, I once again return to the reduced form model since it is much more tractable.

2.1 Setup of the reduced form model

The model has three time periods: $t = 0, 1, 2$. I call these the initial, the interim and the final period. There are two market participants: investors and entrepreneurs, both of them assumed to be risk-neutral. Only entrepreneurs can invest in risky real assets. Investors can only invest in bank debt and require an expected return R_f in the long run (from the initial to the final period). I discuss these and other model assumptions after setting up the full microfounded model, in Subsection 3.1.7.

2.1.1 Markets and banks

There are n markets on a circle indexed by $i = 1...n$, where $n > 3$. There are also n entrepreneurs who can each set up one bank in one of the n markets. I assume no bank can operate in multiple markets. One can interpret this as gains to specialization in a geographical sense or in the kind of activity, e.g. mortgage lending, pursued by the bank.

In market i, bank i can use its expertise to invest a unit at $t = 0$ in a long-term real asset (maturing in the final period) which yields a return of $R_i = R + \epsilon_i - \epsilon_{i+1}$ in the long run, where $R \geq R_f$ is a constant, and $\epsilon_i \sim N(0, \sigma)$ are normally distributed independent random variables with variance σ. Note that ϵ risks in neighboring markets exactly offset each other, thus they can be perfectly hedged away (see Fig. 1).

Entrepreneurs, by establishing a bank, become both managers and equityholders of the corresponding bank, i.e. there are no agency issues between the owners and the management of the bank. As equityholders, entrepreneurs have limited liability. Entrepreneurs have an outside option to invest in an asset with expected return of $R_e > R_f$ in the long run (from $t = 0$ to 2), which they cannot liquidate in the interim period. This can be interpreted as the talent of entrepreneurs, e.g. financial sophistication, being scarce. To achieve this desired higher return, they borrow from investors using debt contracts. Note, however, that entrepreneurs do not have any superior investment possibility in the interim period.

The other type of participant is a continuum of investors, who cannot invest in bank equity, only debt issued by the bank. Investors can be interpreted as uninformed life-cycle savers. Thus bank i is set up using equity (stock) $S_i \geq 0$ provided by entrepreneurs and debt $D_i \geq 0$ provided by investors, such that the total cost of the unit real project held by the bank is financed by either of these two. Thus given that there is no other source of financing, the accounting identity becomes $1 = S_i + D_i$.

Both market participants are risk neutral, have abundant capital at $t = 0$, and value payoff only in the long run. Neither of them gets additional endowment at $t = 1$. This means that the choice of low equity does not come from entrepreneurs not having enough capital ex ante but it is an endogenous choice. All participants in the market have full understanding of the model of the economy and act rationally.

2.1.2 Capital requirement and payoffs

The risk-adjusted capital requirement is computed by a Value-at-Risk measure. If the bank holds a normally distributed risk with standard deviation σ_1 at $t = 1$, it needs to hold a minimum amount of initial capital:

$$\underline{S} = \kappa \cdot \sigma_1$$

e.g. $\kappa \approx 1.96$ gives a 5% VaR. This capital requirement could be regulatory or could be endogenous and enforced by the market. If the bank does not have enough capital in $t = 1$, it goes bankrupt. In the full model, in Section 3, I endogenize the VaR constraint: investors rationally run on banks which do not have enough equity, thus the debt of these banks is not rolled over and they go bankrupt.

If a bank goes bankrupt, it is liquidated for $L < R$. For simplicity I assume that all the liquidation value accrues to the bondholders and cannot be reinvested. Thus if a bank goes bankrupt at $t = 1$ with probability q, holds initial equity of S and borrows at the interest rate of R_l, then the expected payoff of the equityholder is:

$$E[P] = (1-q) \cdot [R - (1-S)R_l - SR_e] + q \cdot (-SR_e)$$

Using that the investors required return is R_f, i.e.

$$(1-S)R_f = (1-q) \cdot (1-S)R_l + q \cdot L$$

after substituting out R_l, one arrives at the following expected payoff to the equityholder:

$$E[P] = R - q \cdot (R - L) - (1-S)R_f - SR_e$$

which is the expression I use throughout when using the reduced form model. Note that through the interest rates on bank debt, the equityholder endogenizes the loss due to liquidation.

2.1.3 Uninsurable idiosyncratic shocks

There is an additional shock which reduces the equity of the bank: I denote the capital reduction of bank i as ξ_i. With probability p the unobservable state of nature is "crisis": one of the banks on the circle receives an uninsurable idiosyncratic shock, its equity capital is reduced by $\xi_j = d$ at $t = 1$. I refer to the state where there is no idiosyncratic shock, the state realized with probability $1 - p$, as "normal" times. Formally we can write:

$$\xi \sim \begin{cases} \text{w.p.} \quad 1-p: & \forall i: \xi_i = 0 \\ \text{w.p.} \quad p: & \forall i \neq j: \xi_j = 0 \\ & \text{for } i = j: \xi_i = d \end{cases}$$

where, in the crisis state, bank j is selected with an equal probability $\frac{p}{n}$ from $\{1, 2...n\}$.

I model the effect of the shock in the following way: the expected value of the real investment drops from R to $R - d$, and for simplicity I assume that this has to be written off in balance sheet equity right away.[7] I assume $L > R - d$, thus it is socially efficient to liquidate banks that get the equity shock.[8] This idiosyncratic shock in the times of crisis is the exogenous shock that can potentially drive the financial system into a systemic crisis. Note, however, that the state of nature being crisis does not lead to there being a systemic crisis in general, that depends on the endogenous equity choice of the banks. In the full model in Section 3, the idiosyncratic shock is uninsurable because of moral hazard reasons.

[7] This assumption is by no means important, I simply assume it for analytical simplicity

[8] This also means that d is large enough that it is not worth to hold extra equity to self-insure against it: I show this later in Subsection 2.2.

2.1.4 Hedging contracts

Since equity is costly, banks take decisions to credibly reduce risk: they hedge their ϵ risks. Recall that the ϵ risks of neighboring banks are exactly offsetting: while bank i holds $+\epsilon_i$, bank $i+1$ holds $-\epsilon_i$. I assume that only banks i and $i+1$ can observe and contract on ϵ_i, that is while they can sign a bilateral contract with each other to hedge ϵ_i, they cannot hedge it with any other bank. One interpretation is that you need to be present and have expertise in markets i or $i+1$ to be able to contract on ϵ_i. Thus the banks hold risks that can be hedged using bilateral contracts, leading to a financial system entangled in bilateral contracts. I assume that the existence of these hedging contracts is observed by all participants.

Even though banks cannot hedge against their own idiosyncratic ξ_i shocks because of moral hazard considerations, they can in effect insure against the idiosyncratic shock of its neighbors, with whom they do business, by signing default insurance contracts on them. I allow for such contracts as well.

2.2 Systemic crisis in equilibrium

I refer to financial networks where the collapse of a single market leads to the collapse of the whole system as contagious. If in crisis all banks fail, I call that a systemic crisis. In this subsection I show that even if the network is contagious, and market participants know that, none of the banks hold enough reserves, since it is a public good with a free riding problem.

Based on the capital requirement, the minimum amount of equity needed to survive, e.g. roll over debt, at $t = 1$ depends on whether the bank has hedged its risks. A bank in autarchy needs to hold equity of at least $\underline{S}^a = \kappa\sqrt{2}\sigma$ (since it holds two uncorrelated normal risks with standard deviation of σ). A bank with only one hedge needs $\underline{S}^u = \kappa\sigma$, while a bank with both risks hedged $\underline{S}^h = 0$. Note that if the system is contagious, then a profitable deviation might be to completely self-insure by holding $\underline{S}^a$ in order to survive in case of crisis. In this case the bank only fails if it is hit directly by the exogenous shock. However, it has to maintain a higher level of equity and the cost of that is higher than that of borrowing. Autarchy is a profitable deviation in case the payoff holding equity $\underline{S}^a$ is higher than the payoff when holding no equity in a contagious system.

$$R - p \cdot (R - L) - 1 \cdot R_f < R - \frac{p}{n} \cdot (R - L) - (1 - \underline{S}^a) \cdot R_f - \underline{S}^a \cdot R_e$$

Note that when holding equity $S = 0$, the full amount of the investment has to be borrowed and the bank has to pay an expected return of R_f on this debt. Thus at least one bank will choose to self-insure by holding substantial equity instead of hedging if:

$$p > p^a = \frac{n}{n-1}\sqrt{2}\kappa\sigma\frac{R_e - R_f}{R - L}$$

The other option, which turns out to be superior to self-insurance, is if the banks choose to set up an insurance pool to increase the equity of the two neighbors of the failing bank, this insurance pool needs funds of $2\underline{S}^u = 2\kappa\sigma$. The intuition behind such a collective bail-out is depicted in Figure 2. The private cost for every participant is to pay the costs of holding $\frac{2}{n}\underline{S}^u$ amount of equity reserves. Such an insurance scheme is sustainable if no bank chooses to opt out. If a bank opts out, the probability that this given bank collapses increases from $\frac{p}{n}$ to $\frac{3p}{n}$, since, given that it is uninsured, it also collapses if either of its neighbors collapses. On the other hand, it does not have to hold any equity at all.

$$R - \frac{3p}{n}(R-L) - R_f < R - \frac{p}{n}(R-L) - R_f - 2\frac{\kappa\sigma}{n}(R_e - R_f)$$

Note that given that this surplus equity is held in the insurance fund and in the crisis state will be transferred to the balance sheet of another bank, all banks still have to borrow the full amount of the investment; they do not hold any equity on their own. Thus the insurance scheme is only sustainable for

$$p > p^i_{out} = \kappa\sigma \cdot \frac{R_e - R_f}{R - L}$$

Clearly $p^i_{out} < p^a$ thus the insurance scheme is already sustainable for crisis probabilities where autarchy is not yet worth it. Thus autarchy does not have to be considered in welfare calculations.

However, the banks do not internalize their effect on the stability of other banks when choosing their equity levels. The insurance is socially optimal if the total welfare of the system with counterparty insurance is higher than that without it:

$$nR - p\cdot(R-L) - n\cdot R_f - 2\kappa\sigma\cdot(R_e - R_f) > nR - n\cdot p\cdot(R-L) - n\cdot R_f$$

where the left hand side is the payoff of all n banks in a stable system with counterparty insurance and the right hand side is that without insurance, in a contagious system. Note that the lenders of the bank get their expected return of R_f irrespective of the choice of banks thus they do not have to included in the welfare calculation. Thus the cutoff value for the crisis probability where counterparty insurance is socially optimal, is:

$$p > p^s = \frac{2}{n-1}\kappa\sigma \cdot \frac{R_e - R_f}{R - L}$$

Clearly $p^s < p^i_{out}$, thus there is inefficiency. In the region (p^s, p^i_{out}), counterparty insurance is not sustainable in equilibrium, even though it is socially optimal. The inefficiency is akin to that in overlapping generations (OLG) models, it is due to pairwise segmented constraints. The constraints in an OLG economy are budget constraints segmented in time, while here they are capital constraints segmented between institutions.[9] The idiosyncratic shock leads to a systemic crisis in equilibrium, given the low choice of equity of the banks. In this interval, mandatory counterparty insurance would be welfare improving, see Section 3.5.1 for discussion of potential policy interventions in the context of the full microfounded model.

[9] I thank Nobuhiro Kiyotaki for pointing this out me.

Up to now, I simply assumed that the idiosyncratic shock d is large enough that the bank does not self-insure to make sure it survives the idiosyncratic shock to equity. Here I show that the socially efficient choice for the given bank, liquidation in case of being hit by the shock, is also the privately optimal one. If the bank wants to be able to survive the idiosyncratic shock, it must hold equity $S = d$, instead of $S = 0$. Thus the payoff without self-insuring must be higher than that with self insuring:

$$R - \frac{p}{n} \cdot (R - L) - 1 \cdot R_f > R - \frac{p}{n} \cdot (R - d) - (1 - d)R_f - d \cdot R_e$$

which always holds if $L > R - d$, which is the assumption we made about d. Note that I assume the idiosyncratic shock is uninsurable e.g. for moral hazard reasons, thus there is no need to check for the viability of insurance.

2.3 Is there underinsurance in practice?

It is important to know whether such a contagious equilibrium is possible in reality. Given that there is no reliable data on bilateral hedges and insurance contracts of the participants in the financial system, I rely on the statements of regulators and market participants to conclude that such an equilibrium is indeed plausible. The potential existence of such a contagious equilibrium was cited as one of the reasons to save LTCM in 1998. The central mechanism modeled in this paper, that losing hedges may induce large losses, was mentioned in the report to the Vice President[10]:

> "Like its other counterparties, the LTCM Fund's OTC derivatives counterparties would have had to re-balance their portfolios in an effort to reduce risk brought on by a default of the Fund. All of these counterparties would have needed to re-establish positions and hedges related to any contracts upon which the LTCM Fund had defaulted. The cost of closing out these positions might have proved greater than the value of the collateral ultimately realized. The risk of loss would have been particularly high for derivatives counterparties of the Fund who were exposed to illiquid risk positions that would have been even more difficult to hedge or liquidate last September."

The idea that derivative contracts make financial markets instable has also been emphasized by Warren Buffett. He once argued that derivatives could lead to a chain reaction:[11]

> "firms that are fundamentally solid can become troubled simply because of the travails of other firms further down the chain. When a "chain reaction" threat exists within an industry, it pays to minimize links of any kind. [...] derivatives are financial weapons of mass destruction, carrying dangers that, while now latent, are potentially lethal."

[10] "Hedge Funds, Leverage and the Lessons of Long-Term Capital Management", Report of The President's Working Group on Financial Markets, April 1999

[11] Berkshire Hathaway Inc., 2002 Annual Report

The basic insight of the paper is that contrary to Warren Buffett's argument, it might not be privately optimal for the banks to minimize links to other banks. To the contrary: banks are willing to link themselves in a contagious system.

3 Full microfounded model

3.1 Model setup

The basic setup of the model (markets and participants) is the same as that described in Section 2. Here I only described the differences. Given the complexity of the full microfounded model, the results in this section are presented in the form of Lemmas and Propositions.

3.1.1 Investments

In addition to the real asset and bank debt, there is a third asset, an outside liquid riskless asset, which I call government bonds. Both agents can invest in government bonds, which have a return of $R_{f,0}$ in the short run (from 0 to 1). For analytical simplicity I set $R_{f,0} = 1$. The return on government bonds from $t = 1$ to 2 is $R_{f,1}$. Government bonds are short-term investments and can be liquidated costlessly at $t = 1$.

The investment technology is slightly more complicated than in the reduced form model. In market i, bank i can use its expertise to invest a unit in a long-term real asset which yields a return of $R_i = R + \epsilon_i - \epsilon_{i+1} - \delta_i$ in the long run, where $R \geq R_{f,1}$ is a constant. The new term, δ_i is an idiosyncratic shock, the realization of which depends on bank i's effort (see Subsection 3.1.2). However, if the investment is liquidated early at $t = 1$, it only yields $L < R$. This idiosyncratic shock to the investment replaces the direct shock to bank capital in the reduced form model. One can think of ϵ's as sector-specific hedgeable risks. The opportunity cost of entrepreneurial capital is $R_e > R_{f,1}$. The amount of government bonds, held by the bank i is denoted by G_i, thus the accounting identity becomes $1 + G_i = S_i + D_i$.

Projects also have a non-pledgable payoff of X: if a bank survives the last period in the model, i.e. it can settle all its contractual obligations, it gets this additional payoff. It can be thought of as reputation, expertise of the bankers, growth opportunities, certain types of intangible capital, etc. It cannot be seized by the creditor in case of default since it is only valuable within the firm, if the firm survives. Even though banks cannot directly borrow against it, it still has an important role in the decision of the banks in the interim period. The idea of using the threat of termination to induce effort is similar to that of Bolton and Scharfstein (1990).

3.1.2 Moral hazard and idiosyncratic shocks

Banks are subject to moral hazard: the idiosyncratic component δ_i of the long-term project of bank i depends on the bank's unobservable effort choice $e_{i,0} \in \{0, 1\}$ at $t = 0$ and $e_{i,1} \in \{0, 1\}$ at $t = 1$.

Banks get private benefit $B_i = B_0 * (1 - e_{i,0}) + B_1 * (1 - e_{i,1})$ depending on their efforts. That is, if they exert full effort in both periods, they get 0, while if they shirk in both periods, they get $B_0 + B_1$. Since the shareholders and the management of the bank are the same, this can be both thought of as direct payoffs to the equityholders or higher wages and perks to the management. For a detailed discussion of the assumption of private benefits, see Subsection 3.1.7. If a bank goes bankrupt at $t = 1$, then $e_{i,1} = 1$ is assumed, i.e. there is only private benefit payoff based on the effort at $t = 0$.

The idiosyncratic shock does not only depend on effort. With probability p the unobservable state of nature is “crisis”: one bank on the circle gets the worst possible idiosyncratic shock, $\delta_j = d$, irrespective of its effort. I refer to the state where idiosyncratic shocks only depend on effort, the state realized with probability $1-p$, as “normal” times. This idiosyncratic shock in the times of crisis is the exogenous shock that can potentially drive the financial system into a systemic crisis, note, however, that the state of nature being crisis does not lead to there being a systemic crisis in general. Formally we can write:

$$\delta_i = max\left[d(1 - e_{i,0}), d(1 - e_{i,1}), \xi_i\right]$$

for all $i = 1...n$. The component of the idiosyncratic shock that is independent of the effort, which I call adverse shock, is:

$$\xi \sim \begin{cases} \text{w.p.} \quad 1 - p: & \forall i : \xi_i = 0 \\ \text{w.p.} \quad p: & \forall i \neq j : \xi_j = 0 \\ & \text{for } i = j : \xi_i = d \end{cases}$$

where, in the crisis state, bank j is selected with an equal probability $\frac{p}{n}$ from $\{1, 2...n\}$.

Thus the idiosyncratic shock can take on a bad value for three reasons: shirking in the initial period, shirking in the interim period or simply bad luck. In normal times, i.e. with probability $1-p$, $\delta_i = 0$ if bank i exerts full effort, and $\delta_i = d$ if it shirks. However, in times of crisis, i.e. with probability p, one of the banks receives $\delta_i = d$ irrespective of its effort, while the other banks get the idiosyncratic shock according to their efforts. This large shock to one bank can be thought of as the possibility of a risk outside the standard model used by market participants or simply a mistake: something that cannot be hedged, but still market participants know there is some small probability of it happening.

Even though market participants do not directly observe efforts, they do get a signal about the expected realization of δ_i at $t = 1$. The signal is:

$$s_i = max\left[d(1 - e_{i,0}), \xi_i\right]$$

That is they get a bad signal about a bank if it either shirked in the initial period or if it was hit by the idiosyncratic shock ξ. Since the effort choices are only observable by bank i, outsiders cannot tell apart bad luck from shirking. Note that the noise in observing $e_{i,0}$ is relatively small, since the only case in which s_i does not fully reveal the effort choice $e_{i,0}$ in the initial period, is the small probability

$\frac{p}{n}$ that bank i is directly hit by the adverse shock ξ_i in crisis. In equilibrium, bad realizations of s_i are due to bad luck but any insurance leads to a serious moral hazard problem, making these idiosyncratic risks non-hedgeable: if it were to be insured, the bank may not have an incentive to exert effort any more. Given that in the equilibrium of this model the banks choose full effort, the idiosyncratic shock driving the results of the model and leading to a potentially systemic crisis is the shock unrelated to effort hitting a single market in case of crisis.

The following important assumptions are made about the moral hazard parameters to make the model relevant to modeling counterparty risk in a contagious network.

Assumption 1. $L > R - d > 0$

This assumption ensures that it is rational for debtholders to liquidate a project when the bank does not exert effort. This is crucial, since otherwise one would need some kind of bank-run framework to ensure contagion. While I do believe this is a possible channel, the main goal of the paper is to show that systemic crisis is possible in an entangled financial system, even without coordination issues.

Assumption 2. $d > B_0 + B_1$

This means that it is socially optimal to exert effort. It is also a very important assumption, since otherwise it would be socially optimal to let the banks shirk, which is not the way one usually thinks about shirking.

3.1.3 Contracts

When issuing debt, banks can choose between long-term and short-term debt financing. By short-term debt I mean that the debt has to be rolled over at $t = 1$, thus debtholders have an option to withdraw funding and force the bank to liquidate its real project. They can also reset the interest rates in the interim period based on the observed behavior of the bank, even if writing such a state contingent contract is hard ex ante. The benefit of short-term debt is that it has a lower interest rate since it can be withdrawn in a state-contingent way and thus can help overcome the moral hazard and risk shifting problem at $t = 0$. This is the rationale behind short-term debt emphasized by Calomiris and Kahn (1991). The benefit of long-term debt, on the other hand, is that it is a safe financing source even in crisis. I do not allow for the renegotiation of debt contracts, a question discussed in Subsection 3.2.7.

As in the reduced form model, banks can write ϵ hedging contracts. Why do banks chose to hedge at all, since debt contracts usually induce risk-shifting? The riskier the bank, the more equity it has to keep in order to receive debt financing from the investors. Since equity is costly, banks take decisions to credibly reduce risk. They choose short-term debt contracts, so in the interim period debtholders can verify the existence of their hedging contracts. Note that the setup is symmetric, so assuming equal bargaining power, the banks do not pay each other for the hedging contract, it has price zero at $t = 0$. Since no information is revealed about ϵ's in the interim period, the price of the contract remains zero at $t = 1$.

Bondholders have seniority to equityholders and get an equal share of the proceeds from liquidating the bank. Furthermore, if the bank defaults in the interim period, all its state contingent hedging and insurance contracts are canceled without payment. However, if banks survive until the final period, then they first have to settle their ϵ hedging contracts before paying back debt.

3.1.4 Timeline of events

The timeline of events is the following. At $t = 0$, in the initial period, banks are set up using debt and equity. They can choose whether to use long-term or short-term debt. Equity is chosen once debt has been raised, thus the debt and the interest rate on it cannot be contingent on any given level of equity.[12] Bank i then decides on whether to invest one unit in real asset i and how much liquid reserves to hold in the form of government bonds. It then chooses effort level $e_{i,0} \in \{0, 1\}$. At the end of the period, banks decide whether to sign bilateral contracts to hedge their ϵ risks.

At $t = 1$, the interim period, all market participants get the signals s_i for all i's. They also get information about all the hedging and insurance contracts outstanding in the market. If investors hold short-term debt they can decide whether to roll over the debt or collect proceeds and invest in government bonds instead. If the bank goes bankrupt due to financing disappearing, it has to liquidate its real project. If bank i receives continued financing, it then chooses effort $e_{i,1} \in \{0, 1\}$.

At $t = 2$, the final period, ϵ's are realized, real investment i yields payoff of R_i, hedging contracts are settled. Banks pay back their debt and all banks that could fulfill all their obligations, i.e. survive, get an extra reputational payoff of X. Finally, all participants consume their payoffs.

Note that the basic timing assumption of information is that the actions of the bank can only be observed ex post, with a delay, i.e. they cannot commit to prudent action ex ante.

3.1.5 Entrepreneur's choice

The entrepreneur's problem is to choose the maximum expected payout, given some constraints. The first choice is between long and short-term debt. This decision is made given the schedule of rates posted by investors. While long-term debt is safer since it does not allow for the early withdrawal of financing, it is also hard to get since the bank cannot credibly commit to exert effort and hedge its risks: in equilibrium long-term debt is not offered by investors. Given that in Subsection 3.2.5 I show that long-term debt financing is impossible with the commitment problems in the model, here I only present the choice in case of short-term debt. A more general approach with debt covenants, where long-term debt is feasible, is discussed in Subsection 3.6.

In case of short-term debt financing, entrepreneur i's problem is to choose debt (bond) D_i, equity (stock) S_i, liquid reserves G_i, effort levels, and hedging contracts to maximize:

$$(1-\pi_1)E_{t=0}\left[max(0, R_i + \epsilon_i - \epsilon_{i+1} - \delta_i + R_{f,1}G_i - R_{l,1}R_{l,0}D_i + C_i)|\text{survive at } t = 1\right]+(1-\pi_1-\pi_2)X-R_eS_i$$

[12] I argue in Subsection 3.6 that the results hold qualitatively even if we relax this assumption.

s.t.

$$(1-\pi_1)E_{t=0}\left[max(0, R_i + \epsilon_i - \epsilon_{i+1} - \delta_i + R_{f,1}G_i - R_{l,1}R_{l,0}D_i + C_i)|\text{survive at } t = 1\right] + (1-\pi_1-\pi_2)X - R_e S_i \geq 0$$

$$1 + G_i = S_i + D_i$$

$$E_{t=1}\left[max(0, R_i + \epsilon_i - \epsilon_{i+1} - \delta_i + R_{f,1}G_i - R_{l,1}R_{l,0}D_i + C_i)\right] + (1-\pi_2)X \geq B_1$$

where C_i is the net payoff from hedging and insurance contracts that bank i has signed. π_1 is the endogenous probability that the bank goes bankrupt at $t = 1$, and π_2 at $t = 2$. Basically to receive the final payoff, both pledgable and non-pledgable, the bank has to survive both periods $t = 1$ and $t = 2$ without ending up bankrupt. Note that the expectation operator is conditional on survival at $t = 1$. The max operator makes sure that payoff at $t = 2$ is only taken into account if the bank can repay all its obligations. It is assumed that in case of bankruptcy, shareholders are wiped out completely, an assumption that is proved in Lemma 3. The first constraint is the participation constraint of the entrepreneur, the second one is the accounting identity that assets equal liabilities. The third constraint ensures that the bank exerts effort at $t = 1$ in normal times. The incentive compatibility constraint at $t = 0$ is not spelled out since if it is not satisfied, investors will not lend in the first place.

3.1.6 Debtholder's choice

Interest rates are determined endogenously, where investors anticipate the banks' equity and investment choices, such that investors break even. The expected return on bank debt has to be the same as the return on government bonds with corresponding maturity. If the bank does not go bankrupt, creditors are repaid fully. In case there is a shortfall of proceeds compared to debt obligations, all creditors receive the same recovery rate. Note that creditors equally splitting the proceeds means that there is no strategic interaction between them, thus liquidation only happens when it is optimal for all bondholders of that bank, i.e. there is no front-running.

I assume that debt does not carry any covenants, thus equity is chosen after debt has been raised. This means that investors cannot make the interest rate dependent on the amount of equity, except when they decide whether to roll over debt. This is an important simplifying assumption, since equityholders can take interest rates as given when making decisions. Also, this assumption basically rules out long-term borrowing, since with long-term debt there is no interim decision to verify a bank's level of equity and risk exposure. However, this assumption is by no means central to the results, Subsection 3.6 discusses the possibility of debt covenants.

Since debtholders are completely rational, i.e. they know the whole system could collapse in crisis, they charge a higher interest rate to offset losses in crisis. Given these interest rates, set by bondholders both on long and short-term borrowing, the banks make their decision on how much to borrow. Thus investors basically offer a schedule of rates on short and long-term borrowing at which banks can borrow any amount. This gives another way of interpreting the absence of covenants: since investors are small, they cannot monitor how much the banks borrowed.

In general, even if the model has a good equilibrium there is a bank-run equilibrium. I assume there is a government or some other planner who has policy tools to avoid inefficient bank runs, thus I rule out the bank-run equilibrium in case there is a non bank-run equilibrium as well.

3.1.7 Discussion of model assumptions

One could argue that the private benefit framework does not readily apply to big financial institutions and one should use risk-shifting instead. There are several reasons for using private benefits in the model. First, debt contracts are in general not optimal for a settings with risk-shifting (Biais et al. 2007), while they are optimal in case of unobservable effort that results in a shift in the distribution of the returns (Innes 1990), as is the case in this model. I do not explicitly deal with the optimality of debt contracts since this is beyond my paper, but the risk shifting framework would definitely undermine the case for debt contracts in general. Second, if one looks at recent years, there is evidence for banks engaging in activities that can be modeled by private benefits. Private benefits can be interpreted as profits that are pumped out of a company before the long-term projects mature, e.g. by using marking-to-model accounting practices that overvalue long-term assets, thus allowing for higher accounting profits. Another way to interpret private benefits is by overpaying employees, since in my model the entrepreneur both owns and runs the bank. Third, one can think of private benefits as an analytically more tractable way of handling risk-shifting if one restricts the contracts to be equity or debt, which are the most prevalent contracts in practice for financing banks and other levered financial institutions. In this latter interpretation, private benefit is the expected payoff to the equity holder from risk-shifting.

A central assumption implicit in the structure of the model is that banks cannot rehedge with other banks if their counterparty fails. In the model this simply follows from the assumption that only two banks can contract on each ϵ_i, thus if a bank's counterparty fails, there is no one else to contract with. Even though this assumption is extreme, it captures the fact that rehedging is very costly in crisis. Evidence from the aftermath of the Lehman default suggests that while many counterparties could rehedge within a trading day with other counterparties[13], they incurred huge losses[14], since they had to rehedge in a very volatile market with counterparties who were less willing to hedge away their risks.

A theoretical question is how to enforce contracts if only two parties can contract on it. The way to think about only two counterparties being able to contract on a given ϵ is that only they are the ones who fully understand it and know the distribution, such that they can write a contract on it and price it. This does not contradict the fact that ex post all banks and even courts can observe the

[13] "CDS market copes with Lehman without clearing house", The Economic Times, September 18, 2008, available online at: http://economictimes.indiatimes.com/articleshow/msid-3495724,flstry-1.cms

[14] "The great untangling", The Economist, November 8, 2008

outcome of ϵ and thus help enforce contracts. Banks who know the underlying distribution can always make a transformation of ϵ to arrive at any distribution skewed in their favor. Thus if a counterparty does not know the distribution of the underlying ϵ , it will not be willing to contract on it, since it is afraid of "being taken for a ride". Note that this happens even though ϵ is observable by all agents in the end.

The circle structure is just imposed for analytical tractability, however, it is important that banks do not have similarly large exposures to all other banks in the system. The assumed risk distribution is not unrealistic given that most risks can be hedged away with a few contracts or counterparties. The general equilibrium setup of Section 4 proves this intuition in the case where the network structure is endogenous. Generally banks tend not to have a similar exposure to all banks in the whole world, there are always a few counterparties that are the most important. Having only two natural counterparties captures this insight.

In the baseline model, debt does not have any covenants, thus debtholders cannot force the bank to hold a given level of equity or to sign a given contract. This assumption is used for two reasons: one is that debtholders, especially in form of bonds, indeed have limited influence on the financing and investment choices of a large financial institution. Given that these institutions don't usually have large creditors and that recently (at least before 2008) most bonds have been covenant-lite, this is a reasonable assumption. The other reason is simplicity: one can then ignore the effect of the firm's decisions on the interest rate it is charged by debtholders in the initial period. On the other hand one could worry whether the same results hold if debt covenants are allowed.[15]

Given that the model is highly stylized, there can be multiple interpretations of the ϵ risk. One can think of ϵ as a bilateral or OTC contract, such as a credit default swap on a third party. However, it is more general than that: it can be an ongoing business relationship, a contractual guarantee or a state-contingent credit line.

The non-pledgable payoff X is a crucial element of the model. The main effect driving the results is exactly that the expected value of the non-pledgable payoff decreases if a bank becomes risky, since there is a higher probability of bankruptcy. Note that in general, equityholders, holding a convex payoff function, gain from increased risk, but to model contagion they must, on the contrary, lose from higher risk. In this model with short-term debt, the interest rate is reset to offset any gains to equity from increased risk: all losses are due to the decrease in expected value of the non-pledgable payoff. The two effects result in an overall decrease of equityholder value if the bank becomes risky, thus leading to the violation of the incentive constraint in crisis, which is in turn the main source of contagion.

The model assumes that there is only one bank in every market. This bank can be thought of as a representative bank in a given market, the model generalizes to the case with multiple banks following the same strategy in a market. In this latter case, the idiosyncratic shock is an unexpected drop in value of a market, not a single bank. Thus I use the terms bank and market interchangeably.

[15]In Subsection 3.6 I discuss this possibility and argue that the non-contractibility of effort is enough to yield inefficiency.

In case one would like to analyze the effect of the failure of a single bank with multiple banks in the same market, we would have to specify the nature of competition between the banks: this analysis, although potentially interesting, is beyond the scope of the current paper.

3.2 Systemic crisis in equilibrium

3.2.1 Equilibrium with contagion

Under certain parameter restrictions, I prove the existence of the following subgame perfect Nash equilibrium. Banks choose to finance their operations using short-term debt, which is used as an incentive device to ensure prudent behavior at $t = 0$. Since equity is costly, banks choose not to hold more equity than what is just enough to make the incentive constraint at $t = 1$, to exert effort in normal times, hold. This means there is no slack equity, i.e. equity beyond what is needed in normal times, to use in crisis: so when one of the banks defaults, all banks default.

These consecutive (in the model, contemporaneous) defaults are driven by short-term debtholders withdrawing their financing from risky banks one after the other. Banks become risky one after another since they lose their ϵ hedges when the neighboring bank defaults. Once a bank has lost one of its hedges, it becomes risky, thus the probability that it ends up bankrupt at $t = 2$, because it cannot repay its creditors, increases. This increase in the probability of default in turn reduces the expected value (as of $t = 1$) of the non-pledgable payoff. Since the non-pledgable payoff is used as reputational collateral in the incentive constraint, the incentive constraint is now violated. Thus it is optimal for debtholders to withdraw their funding and liquidate the project instead of letting it mature while the bank does not exert effort.

In the following I show, step by step, that the proposed equilibrium exists: this yields the micro-foundation for the reduced form model of Section 2. First, we conjecture that banks use short-term debt and hedge their ϵ risks. Lemma 2 calculates how much is the minimal equity needed to ensure that debt is rolled over in the interim stage given the level of riskiness of the bank's assets. Lemmas 3 and 4 prove that the contagion mechanism indeed leads to the collapse of all connected banks. Then, in Lemmas 5 and 6, we verify that indeed short-term debt financing is chosen over long-term debt and that hedging of ϵ risks is optimal for banks. Finally, Proposition 1 shows that if the probability of crisis is small, then banks endogenously choose to hold only the minimum amount of equity and thus the system is contagious and completely collapses in crisis.

3.2.2 Entrepreneur's problem with short-term debt and ϵ hedging

I first show in Lemma 1 that holding government bonds is suboptimal for the banks.

Lemma 1. *Banks optimally do not hold liquid assets: $G_i = 0$ for all i.*

The proof of this lemma is relegated to the Appendix along with all other proofs omitted in the text. From now on I set $G = 0$, since by Lemma 1 holding government bonds is never optimal for a bank. The intuition is as follows: simply increasing the reserves held in government bonds G, keeping the amount of equity at the same level, does not alleviate the incentive constraint. Since the increase in government bond reserves, given fixed equity, is financed by issuing bonds at an interest rate higher than the riskless rate, it is costly for the entrepreneur. Given that it is financed by debt, it cannot increase the probability of survival in crisis either, since it does not increase slack equity. Thus the only way to make sure that the incentive constraint is not violated in crisis, if the bank's neighbor fails, is to hold a higher equity stake ex ante. I refer to this as holding slack equity, since in normal times, this equity is indeed too much, at least ex post.

For now we conjecture that the banks use short-term debt and that they hedge their ϵ risks. I verify this in Lemmas 5 and 6. For simplicity I drop the subscript i since entrepreneurs are ex ante symmetric. Given the conjecture, the entrepreneur's problem now reduces to the following: choose debt D and equity S to maximize

$$[1 - \pi(R_{l,0}, R_{l,1}, D)] \cdot (R - R_{l,1}R_{l,0}D + X) - R_e S$$

s.t.

$$[1 - \pi(R_{l,0}, R_{l,1}, D)] \cdot (R - R_{l,1}R_{l,0}D + X) - R_e S \geq 0$$

$$1 + G = S + D$$

$$R - R_{f,1}R_{l,0}D + [1 - \pi_2(R_{l,0}, R_{l,1}, D)] \cdot X \geq B_1$$

Remember that $\pi(R_{l,0}, R_{l,1}, D)$ is the endogenously determined probability of the bank going bankrupt at either $t = 1$ or $t = 2$ given the asset and liability choice. It can be split into the probability of failure at $t = 1$ of π_1, and that at $t = 2$ of π_2. Note, however, that the probability of default in the interim period π_1 depends not only on the choice of the individual bank but on the choice of all other banks in the financial system. In the contagious equilibrium considered in this subsection, $\pi = \pi_1 + \pi_2 = p + 0 = p$ holds, since the whole network collapses in the case of crisis at the interim period but there is no systemic default in normal times. Note that counterparty insurance contracts are not considered to be part of the choice set of banks for now, these are analyzed in Section 3.3. It is assumed that in case of bankruptcy, shareholders are wiped out completely, a conjecture that is proved in Lemma 3.

The incentive constraint at $t = 1$ was rewritten based on the following observation. The pledgable payoff of the project is R independent of the debt level. The expected payoff of the investors to make them brake even on the debt of face value $R_{l,0}D$ at $t = 1$ is $R_{f,1}R_{l,0}D$. Thus the expected pledgable payoff to the entrepreneur is $R - R_{f,1}R_{l,0}D$. The non-pledgable payoff is conditional on not being bankrupt at $t = 2$ if the bank survived at $t = 1$: $[1 - \pi_2(R_{l,0}, R_{l,1}, D)] \cdot X$. The left hand side of the incentive constraint, the expected total payoff to entrepreneur if he exerts effort at $t = 1$ is the sum of these two terms.

3.2.3 Minimal equity needed for rolling over debt

First we calculate the minimal amount of equity $\underline{S}(R_{l,0}, \sigma_1)$, that needs to be held by the bank to make sure the incentive constraint is satisfied in the interim period if it holds risks of standard deviation σ_1 at $t = 1$. Equity needs to be held since creditors are worried that banks could misbehave in the interim period and collect private benefits B_1. The minimum amount of equity $\underline{S}(R_{l,0}, \sigma_1)$ is thus increasing in B_1, the temptation for banks to shirk. Note that this equity level also increases in $R_{l,0}$ since the higher the interest rate on the initial debt, the more has to be repaid in the end.

The following assumption ensures that banks have to keep at least some equity to overcome moral hazard. Basically it assumes that moral hazard is a serious problem in the financial system. Given that financial institutions cannot typically operate with negative equity, this restriction seems reasonable.

Assumption 3. $B_1 \geq (R - R_{f,1}) + X$

Lemma 2. *Assume a bank has a final payoff with expected value R and normally distributed risk σ_1 at $t = 1$ and is financed by short-term debt. The minimum amount of equity to be held by this bank to avoid a withdrawal of debt financing at $t = 1$ is given by the following implicit equation:*

$$\underline{S}(R_{l,0}, \sigma_1) = \frac{B_1 - (R - R_{f,1} R_{l,0}) - [1 - \underline{\pi}_2(R_{l,0}, \sigma_1)] \cdot X}{R_{f,1} R_{l,0}} > 0$$

where the endogenous probability of default at $t = 2$ is:

$$\underline{\pi}_2(R_{l,0}, \sigma_1) = \Phi\left(\frac{\underline{R}_{l,1} R_{l,0}[1 - \underline{S}(R_{l,0}, \sigma_1)] - R}{\sigma_1}\right)$$

and Φ denotes the cumulative density function of the standard normal distribution. The short-term rate $\underline{R}_{l,1}$ charged to the bank at $t = 1$ is determined endogenously. Also, the minimum amount of equity increases with risk held by the bank:

$$\frac{\partial \underline{S}}{\partial \sigma_1} > 0$$

Proof. The incentive constraint of a bank surviving at $t = 1$ is the following:

$$R - R_{f,1} R_{l,0} \cdot D + (1 - \pi_2(R_{l,0}, \sigma_1)) \cdot X \geq B_1$$

The bank chooses $S = 1 - D$, s.t. the incentive constraint binds: this yields the above implicit expression for the equity. Assumption 3 implies that $\underline{S}(R_{l,0}, \sigma_1) > 0$ holds always, since $R_{l,0} \geq 1$.

Denote the risk held by the bank at $t = 1$ by $\eta \sim N(0, \sigma_1)$. The bank defaults in $t = 2$ if and only if:

$$R + \eta < R_{l,1} R_{l,0} D$$

Given the normal distribution of η, the probability of default at $t = 2$ is as stated in the lemma.

$\underline{R}_{l,1}$ is determined by investors breaking even on the loan of size $R_{l,0}\underline{D} = R_{l,0}(1-\underline{S})$ at $t=1$:

$$R_{f,1}R_{l,0}\underline{D} = \underline{R}_{l,1}R_{l,0}\underline{D} - \int_{-\infty}^{\underline{R}_{l,1}R_{l,0}\underline{D}} (\underline{R}_{l,1}R_{l,0}\underline{D} - x)\frac{1}{\sigma_1}\phi\left(\frac{x-R}{\sigma_1}\right)dx$$

where the expression on the right hand side is the nominal debt payment minus the expected shortfall due to bankruptcy. This simplifies to the following numerically feasible implicit equation for $\underline{R}_{l,1}(R_{l,0},\sigma_1)$:

$$(\underline{R}_{l,1} - R_{f,1})R_{l,0}\underline{D} = \sigma_1 \cdot \phi\left(\frac{\underline{R}_{l,1}R_{l,0}\underline{D} - R}{\sigma_1}\right) - (R - \underline{R}_{l,1}R_{l,0}\underline{D}) \cdot \Phi\left(\frac{\underline{R}_{l,1}R_{l,0}\underline{D} - R}{\sigma_1}\right)$$

For the proof of $\frac{\partial \underline{S}}{\partial \sigma_1} > 0$, see the Appendix. □

The minimal amount of equity $\underline{S}$ is decreasing in the bank's expected non-pledgable payoff $(1-\pi_2)\cdot X$. This is important since it highlights that non-pledgable payoff is crucial: it is used as a kind of reputational collateral to overcome the moral hazard problem. Thus if the expected value of non-pledgable payoff decreases because of the risk of bankruptcy (rise in π_2), more equity is needed to ensure high effort. In fact the increase in the required amount of equity in crisis is the main mechanism that leads to withdrawal of short-term financing.

3.2.4 Minimal equity is not enough in crisis

Denote $\underline{S}^h(R_{l,0}) = \underline{S}(R_{l,0}, 0)$ the amount of equity a bank would hold if it perceived its project as completely riskless because of perfect hedging. To show that all banks collapse in crisis if they all hold equity of only $\underline{S}^h(R_{l,0})$, one has to reason that banks collapse one-by-one if a single market (bank) fails. One-by-one is only meant intuitively since in equilibrium all banks fail contemporaneously at $t=1$. First, in Lemma 3, I show that the bank, the market of which is hit in crisis by the adverse ξ shock, fails. As a next step, in Lemma 4, I show that if bank i collapses and leaves its neighbor (counterparty), bank $i+1$, unhedged, that neighbor will also suffer a run of creditors, i.e. its debt will not be rolled over and thus it will fail too. This in turn leaves the next bank on the circle, bank $i+2$ unhedged and leads to the collapse of that too. If all banks keep only the minimal equity of $\underline{S}^h(R_{l,0})$, this mechanism leads to the "sequential" collapse of all banks on the circle. Since these banks are the only ones who can finance projects in their own markets, all real projects will be prematurely liquidated at a large real cost.

The following assumption is introduced to ensure that the bank does go bankrupt if its debt is not rolled over in the interim period. Since banks usually operate with high leverage, such as 10 for commercial banks and 30 for investments banks before 2008, it is plausible to assume that a withdrawal of debt financing would force them to prematurely liquidate their real assets and even after liquidation they would not be able to meet all debt obligations. The implicit assumption here is that forced liquidation is costly, thus L is substantially smaller than 1.

Assumption 4. $R - B_1 > R_{f,1}L$

Lemma 3. *If a bank only holds equity of $\underline{S}^h(R_{l,0})$, it goes bankrupt if its debt financing is withdrawn at $t = 1$. If for bank i, the signal about final payoff at $t = 1$ is the lowest possible, $s_i = d$, then its creditors do not roll over its debt and thus the bank has to be liquidated because it is bankrupt.*

Proof. A bank will go bankrupt at $t = 1$ if the liquidation value of the project is not enough to cover debt obligations, i.e. $R_{l,0}D > L$. If the minimum level of equity is held, implying debt level of $\underline{D}_h(R_{l,0}) = 1 - \underline{S}^h(R_{l,0})$, the condition for bankruptcy is $R - B_1 + X > R_{f,1}L$, which in turn is satisfied by Assumption 4. For banks hit directly by an adverse idiosyncratic shock, i.e. $s_i = d$, the long-term return on the real project is $R - d$ irrespective of the bank's effort at $t = 1$. Bankruptcy follows from Assumption 1, i.e. that $R - d < L$, and the investors optimally liquidate the real investment. $\square$

The interpretation is straightforward. Look at the bank, the long-term project of which is guaranteed to have a bad payoff, either because of the shock ξ or because of an unhedged risk. It is optimal for its creditors to demand the project be abandoned and liquidated, since that is still better than letting this bad project run. In fact, this is exactly the rational for short-term debt: creditors want to have the right to terminate a project if it is surely going to have a very low return e.g. when the bank did not exert sufficient effort in the initial period. Note that at this point ruling out renegotiation is important. If debtholders could renegotiate debt and let the bank survive if its neighbor failed, the system would not be contagious. See Subsection 3.2.7 for a discussion on why ruling out renegotiation is reasonable.

The next step, formalized in Lemma 4, is at the heart of the contagion mechanism: if all banks hold only an equity of $\underline{S}^h(R_{l,0})$, the failure of a neighbor, with which it hedged an ϵ risk, leads to a violation of its incentive constraint, thus it would choose to shirk so it is also preemptively liquidated by its creditors.

Lemma 4. *In case a bank holds equity of $\underline{S}^h(R_{l,0})$, it exerts full effort at $t = 1$ if both of its neighbors survive. However, it chooses to shirk if one or both of its neighbors defaults leaving it unhedged: its debt is not rolled over at $t = 1$, it goes bankrupt and its real project is liquidated.*

Proof. If neither of the neighbors default, exerting effort is optimal, since the incentive constraint is satisfied when equity of $\underline{S}^h(R_{l,0})$ is chosen. If one or two of the neighbors defaults, the risk of payoff as of $t = 1$, σ_1, increases. By Lemma 2, $\frac{\partial \underline{S}}{\partial \sigma_1} > 0$, so a higher level of equity would be needed to roll over debt. Debt financing is thus not rolled over and the bank goes bankrupt. Note that any change in the interest rate $R_{l,1}$ simply offsets the effects of risk shifting, thus the interest rate charged at $t = 1$ does not change the expected pledgable payoff of the entrepreneur at $t = 1$: the incentive constraint is only effected by the change in the expected value of non-pledgable payoff. $\square$

Lemma 4 highlights the main mechanism of contagion through the loss of hedging contracts. The intuition is as follows: if a bank collapses and leaves its neighbor unhedged, that bank will become more risky. Becoming more risky means that there is a larger probability of it going bankrupt at the final date of $t = 2$. This in turn reduces the expected value of non-pledgable payoff beyond the long run. However, non-pledgable payoff was used as collateral in the incentive problem: basically it is used by banks to be able to commit to high effort even with relatively low levels of equity. This mechanism, to my best knowledge, has not been proposed previously as a potential cause for a cascade of bank runs induced by wholesale creditors or bondholders.

The main point is that equity that is enough in normal times is not enough in times of crisis. Banks of course make their equity decisions based on the probability of crisis. As I show in Proposition 1, if the probability of crisis is low, banks choose to simply hold the minimal amount of equity $\underline{S}^h(R_{l,0})$ and thereby induce contagion in the system: they make the crisis, that was originally idiosyncratic, systemic.

3.2.5 Choice of debt maturity and hedging

Now I show in Lemma 5 and 6 that ϵ-hedging and short-term borrowing is the optimal choice for banks. Given the assumption that banks cannot commit to a capital structure at $t = 0$, long-term borrowing is not feasible. Lemma 5 below formalizes this intuition. However, to ensure that short-term borrowing is feasible, even though long-term borrowing is not, one has to place an upper limit on B_0, the private benefit from shirking in the initial period. Furthermore, to ensure that the entrepreneur's participation condition is satisfied ex ante, one has to place an upper bound on the return on the outside option to entrepreneurs. The following, not too intuitive, Assumption 5 does both. Note that the coefficient of R_e is positive (by Assumption 3), making Assumption 5 an upper bound on R_e.

Assumption 5. $(1-p)B_1 > B_0 + R_e\left[1-(1-p)\frac{R+X-B_1}{R_{f,1}}\right]$

Lemma 5. *Since banks cannot commit to a given equity level at $t = 0$, long-term borrowing is not offered by the investors. Thus entrepreneurs use short-term debt and they are willing to participate in the market.*

Long-term lending is impossible under these assumptions since if the investor has no mechanism to enforce that the bank has some level of equity and contracts at the interim period, there is no way the bank can commit to exerting effort. Once the bank has raised long-term debt it is always worthwhile to shirk and collect private benefits. This result does not hinge on the extreme assumption of no debt covenants, it does hold in case there are debt covenants, see Subsection 3.6.

The next step is to show that, in equilibrium, the banks choose to hedge their ϵ risks with their neighbors.

Lemma 6. *For any $i = 1...n$, bank i completely hedges risk ϵ_i with bank $i+1$, and risk ϵ_{i-1} with bank $i-1$ in case debt is short maturity. These hedging contracts are long-run contracts.*

The basic insight is that with short-term debt, banks are punished by higher borrowing rates, or even withdrawal of funds in the interim period, if they do not hedge their ϵ risks. Thus even though the entrepreneur's payoff is a convex function of the pledgable payoffs of the bank's operations, the entrepreneur cannot gain from shifting risk to debtholders. Furthermore, not hedging also decreases the expected value of their non-pledgable payoff, through the increased probability of bankruptcy, so all in all they lose by not hedging their ϵ risks. The Lemma also highlights the strong incentive to get entangled in these bilateral hedging contracts: hedging decreases the amount of equity that has to be held in order to satisfy incentive constraints, since a hedged bank is less risky.

It is not simply that banks engage in hedging to lower the capital requirement set by the regulator, there is an underlying moral hazard issue that is the "raison d'etre" of these contracts. Thus the results of the microfounded model help us understand that, in general, abolishing VaR type capital constraints do not solve the problem of systemic crisis. The VaR-type constraints arise endogenously from the underlying informational frictions.

3.2.6 Contagious system in equilibrium

A potential deviation from the proposed Nash-equilibrium with short-term debt is that a bank in a contagious banking system chooses to hold a higher amount of equity, to make sure it survives the crisis. Note that it needs enough equity to survive losing both of its ϵ hedges, given the circular structure of the markets. Thus this is also the amount of equity required for unhedged, stand-alone autarkic banks holding total risk of $\sqrt{2}\sigma$. Denote the amount of equity an unhedged bank has to hold by $\underline{S}^a(R_{l,0}) = \underline{S}(R_{l,0}, \sqrt{2}\sigma)$. If an equilibrium exists with all banks holding $\underline{S}^a(R_{l,0})$, then no ϵ hedging contracts are needed between banks to insure that the incentive constraint in the interim period is satisfied.

Thus to conclude that the contagious equilibrium exists, one has to make sure it is not a profitable deviation for any of the banks to hold higher amounts of equity in order to self-insure against the collapse of neighboring banks in a systemic crisis. Intuitively, if the probability of a systemic crisis is high, it is worth to self-insure, while if it is low, given the high costs of holding equity, each bank chooses not to self-insure. Proposition 1 formalizes this intuition and gives a cut-off value in the probability of crisis p. To simplify the analysis, I assume the following lower bound on σ, it is by no means a necessary condition:

Assumption 6.

$$\sigma \geq \frac{\sqrt{2}}{\sqrt{\pi}} \cdot \frac{p \cdot X}{\frac{R_e}{R_{f,1}}\left(1 - \frac{p}{n}L\right) - 1}$$

Proposition 1. *There exists a $p^a > 0$, such that if $p < p^a$, then there is an equilibrium where all banks hold equity of only $\underline{S}^h(R^h_{l,0})$ and the system completely collapses if a single bank fails, i.e. if the financial network is contagious. The implicit equation for p^a is:*

$$p^a = \frac{R_e[\underline{S}^a(R^h_{l,0}) - \underline{S}^h(R^h_{l,0})] - \tilde{\pi}_2 X}{\frac{n-1}{n}B_1 - \tilde{\pi}_2 X}$$

The equation is implicit for the cutoff value of p^a since both $\underline{S}^a(R^h_{l,0})$ and $\underline{S}^a(R^h_{l,0})$ depend on p^a through $R^h_{l,0}$. $\tilde{\pi}_2$ is the endogenous probability of bankruptcy. In the specific case where the system is *contagious and all banks fail in crisis, the endogenous short-term interest rate at* 0 *is:*

$$R^h_{l,0} = \frac{1}{1 - p\left(1 - \frac{R_{f,1}L}{R+X-B_1}\right)}$$

Furthermore, the entrepreneurs' participation constraint holds in this equilibrium.

Note that the above proposition does not assess whether this contagious equilibrium is indeed socially optimal or not. I show in Section 3.3, in a more general setting allowing for counterparty insurance, that this contagious equilibrium is not socially optimal in general.

Notice that the initial interest rate $R^h_{l,0}$ in a contagious system, where all banks hold equity of $\underline{S}^h(R^h_{l,0})$, is set by the investors such that they are completely compensated for the losses when the whole system collapses in contagion. As the probability of crisis p increases, the interest rate also increases (the coefficient of p is positive by Assumption 4).

The aspect of the model that the failure of a single market, which is an idiosyncratic risk, leads to the collapse of all other markets is an important point from a macroeconomic point of view. The model thus gives an insight on how a single idiosyncratic shock, which is a small shock from a macro perspective (given that n is large), can lead to a macroeconomic shock. This mechanism shows how the financial system can generate sizeable "productivity shocks", an issue that has puzzled macroeconomists for decades. While the idea that the financial sector acts as an amplifier is not in itself new, the property of the model to generate a multiplier of any magnitude, is not common in models of the financial system; the model of Kiyotaki and Moore (1997) is one of the few exceptions. These model economies are sometimes called "butterfly" economies, referring to the nature of chaotic systems, that a small disturbance in the atmosphere can cause a hurricane, but are usually modeled by mechanical systems without optimizing agents.[16]

This is a model of an economy where the financial sector and thus real projects do not get enough debt funding in crisis. This is not due to supply effects, there is ample amount of liquidity outside the financial system. The lack of funding is due to the inability of this liquidity to enter the financial system in the interim period because of the incentive constraints of the banks that are violated. In case of crisis, there is a "flight to quality" of investors' funds. Instead of rolling over the debt of financial institutions, they choose to invest in government bonds. Thus the model does give an explanation for these episodes in finance and highlights the important role that incentive constraints play in "flight to quality".

3.2.7 The question of renegotiation

The possibility of renegotiation arises in this setting too, like in many models of corporate finance. However, given that large financial institutions usually issue bonds and commercial papers, renegotiation does not seem to be a viable option in practice, this is why I rule out this possibility in the

[16] For an example of amplification in a mechanical model see Bak et al. (1992).

model. There are two reasons why these institutions issue these very incomplete contracts, which are basically impossible to renegotiate. First, given the vast amount of money they borrow, they have lots of relatively small lenders. Second, most of these lenders are financially not very sophisticated, they prefer simple contracts, such as debt without all kinds of embedded options and state-contingent payments. Even if it is possible to observe the event of a neighbor failing, it might be impossible to write a contract ex ante that fully describes the uncertainty and the amount of state-contingent debt forgiveness, let alone sell it to investors who are not that financially sophisticated. This asymmetry in financial sophistication is exactly what the model tries to capture by having two types of agents: entrepreneurs and investors.

Assume for now that renegotiation or some kind of state-contingency is allowed with debt contracts. Any contract that grants debt forgiveness to the firm that is hit by the adverse idiosyncratic shock δ_i is subject to a moral hazard problem (see Lemma 7 for the proof). However, forgiving some debt to the bank, the neighbor of which failed, is not subject to such a problem, since the default of the neighbor does not depend on the bank's effort. Thus a debt contract that forgives just enough debt to the bank, the neighbor of which defaulted, such that its incentive constraint is no more violated, would be a Pareto improvement. In practice such contracts might not be possible for the reasons mentioned before. Furthermore, it might give banks the incentive to contract with weak counterparties who offer cheap hedging, another source of moral hazard which, however, is not modeled in this paper.

Notice that in general the maturity structure of a financial institution also allows for front running: the shortest term creditors can easily withdraw, regaining most or all of their money, thus they are not interested in renegotiation. He and Xiong (2009) show in a dynamic setting with small lenders, that investors holding the shortest term debt will run if the fundamental of the bank deteriorates or it becomes more risky, even if the bank itself is solvent. This result highlights that increased risk can lead to run and that renegotiation is unlikely in practice. It also makes clear that the modeling choice of this paper that banks have to become insolvent as of $t = 1$ for short-term creditors to run, is not crucial for the results.

3.3 Suboptimal counterparty insurance

3.3.1 The non-insurability of the idiosyncratic shock

Since the adverse idiosyncratic shock ξ in crisis is so devastating to the system, it might be worthwhile to insure against it. However, due to the moral hazard problem, such an insurance is not possible. The following lemma formalizes this argument.

Lemma 7. *No insurance scheme aimed at saving the bank, hit directly by the adverse idiosyncratic shock in crisis, is feasible.*

Note that this also means that any government policy which saves failed banks to prevent systemic crisis creates a moral hazard problem, a strong argument brought forth against the bailout package during the crisis of 2008. Thus relying on ex-post bailouts is not a reasonable policy in this model, a better way is to have insurance schemes that aim at stabilizing the neighbors and business partners of the failing bank. In the following subsection, I explore whether such a counterparty insurance scheme can be implemented by the market itself.

3.3.2 Equilibrium with suboptimal counterparty insurance

The main question of the paper is whether the contagious nature of the financial system is welfare reducing or whether it is simply an optimal outcome, like in the model of Allen and Gale (1998). To assess this question, one has to allow for possible state contingent contracts the banks would like to sign with each-other to potentially diversify counterparty risk, since insuring against the adverse idiosyncratic shock directly is not feasible (see Lemma 7). Under such a scheme, a bank would get a state contingent payment, on the condition that its neighbor with which it was hedging its ϵ risk, fails. This state-contingent payoff can then be used to raise the bank's equity. If this increase in equity is large enough, the bank can survive even if its counterparty defaults, leaving it unhedged, since its incentive constraint is no more violated.

Now assume that a counterparty insurance scheme is in place that allows for diversification of counterparty risk. A bank needs more equity in order to have its debt rolled over in case one of its ϵ hedging counterparty fails in a crisis and it is exposed to risk of standard deviation σ. Denote this level of equity by $\underline{S}^u(R_{l,0}) = \underline{S}(R_{l,0}, \sigma)$. Since the failed bank has two neighbors, both in need of higher equity, excess equity reserves of $2[\underline{S}^u(R_{l,0}) - \underline{S}^h(R_{l,0})]$ are needed to stabilize both neighbors. Note that under such a scheme, the system is stable, so no bank has to worry about both of its counterparties failing. Second, I argue that it is possible to set up an insurance scheme where all other banks help the banks in trouble to raise equity. The basic insight is that the default of a neighbor is an observable event, furthermore it is contractible since it does not involve moral hazard considerations: the bank that could potentially have been shirking is let to fail.

The optimal insurance mechanism that banks can engage in, is the following. They keep just enough slack equity, i.e. equity beyond what is needed in normal times, s.t. if they transfer this to the two troubled banks neighboring the one that failed because of the adverse shock, then both of them can roll over their debt. The basic intuition of the scheme is depicted in Figure 2. There are two ways to store this slack equity. One is to create a central insurance fund that invests in government bonds and pays out to the neighbors of the failed bank in crisis. Banks contribute in an equal amount to the fund since they are symmetric and if no insurance payout is made, the proceeds of the fund are returned to the banks. Note that while here insurance payments are the same, while in a heterogenous financial network, they could differ from bank to bank: this is explored in Subsection 5.

A slightly less efficient but less centralized scheme is to keep the reserves for insurance on the books of the banks, in form of slack equity. Then banks can sign default swaps with all banks except its neighbors, on its neighbors as reference entities. While this allows the market to price the insurance in the form of default swaps, it is much less transparent, furthermore it requires slightly higher slack equity. The amount of slack equity has to be increased to $\frac{n}{n-1}$ times the amount of reserves in the central insurance fund since some of the slack equity will be lost in the bank which is directly hit by the adverse shock. Thus from a social perspective, keeping the reserves in form of slack equity is socially suboptimal compared to having a common pool.

Proposition 2 contains one of the main results of the paper: in a range of crisis probabilities p, counterparty insurance, while socially optimal, cannot exist in a competitive equilibrium. It also shows that the inefficiency result of the simple reduced form model (Subsection 2.2) holds up in the full microfounded model.

Proposition 2. *There exist thresholds p^i and p^s, st. $p^i > p^s > 0$ and*

(i) if $p < p^s$, it is not socially optimal to insure against counterparty risk

(ii) if $p^s \leq p < p^i$, a counterparty insurance scheme is socially optimal but in equilibrium there is no counterparty insurance, and the failure of any bank causes a systemic crisis, i.e. all banks to fail

(iii) if $p \geq p^i$, a voluntary counterparty insurance scheme is sustainable in equilibrium and the crisis is not systemic

If the reserves to cover insurance payments are kept in a separate central fund investing in government bonds, the implicit equations for p^s and p^i are:

$$p^s: \qquad p = \frac{2}{n-1} \cdot \frac{R_e - R_{f,1}}{R_{f,1}\hat{R}_{l,0}} \cdot \frac{\hat{\pi}_2 X}{R + X - R_{f,1}L - \frac{2}{n-1}\hat{\pi}_2 X}$$

$$p^i: \qquad p = \frac{R_e - R_{f,1}}{R_{f,1}\hat{R}_{l,0}} \cdot \frac{\hat{\pi}_2 X}{B_1 - \hat{\pi}_2 X}$$

the equations are implicit since $\hat{R}_{l,0}$, the endogenous interest rate at $t = 0$, depends on p. $\hat{\pi}_2$ is the endogenous probability that a bank goes bankrupt at $t = 2$ if it is the neighbor of a failed bank and got a capital infusion through the insurance scheme.

The main question is where this inefficiency comes from. The source of externality is that banks, when deciding on how much reserves or slack equity to hold in order to avoid bankruptcy, do not internalize their effect on their neighbors through the ϵ hedging contracts in case they go bankrupt. There is a free riding problem too: every bank wants the others to insure against counterparty risk and thus stabilize the system, while they themselves would rather not contribute through costly equity, they prefer to simply enjoy the benefits of a non-contagious financial system.

To formalize the intuition a bit, imagine each bank makes a profit of V if it survives in the interim period. In a non-contagious system, in the initial period, every bank perceives the risk of losing V with probability $\frac{2p}{n}$, i.e. when one of their neighbors defaults. Thus the maximum amount that can be charged to banks as the cost of insurance is $\frac{2p}{n}V$, totalling $2pV$ for the whole system. However, in the case without insurance, the system is contagious, thus all real projects are liquidated and all banks go bankrupt. This means that the social benefit of insurance for the system as a whole is npV. Clearly, for $n > 2$, the social benefits of insurance cannot be covered by voluntary private contributions since $2pV < npV$.

The region of efficiency and inefficiency is illustrated in Figure 3. An important insight is that $\frac{p^i}{p^s}$ is on the order of n, thus the inefficient region is large compared to the region where no counterparty insurance is optimal, furthermore, this ratio grows with the size of the system. This does not come from an increase in the socially optimal cutoff probability p^i, but from a decrease in the privately optimal cutoff probability as the size of the system increases.

Note that as $p \to 0$, the financial system completely collapses in crisis, however, not insuring against the crisis is socially optimal. This case is reminiscent of the analysis of Allen and Gale (2000), even though their mechanism is different, since market participants make decisions without taking the possibility of a crisis into account. It also highlights the weakness of their study, since as $p \to 0$, the crisis is socially optimal: it is not worth to hold large costly reserves to avoid a bad outcome that is very unlikely.

One might wonder whether so called "due diligence" contracts could overcome the problem of inefficiency. These could specify that the counterparty has to hold insurance against the default of its own counterparties. Efficiency is not restored however, since it would still only incorporate the private benefits to the two counterparties, but not to the system as a whole. Counterparties could still decide to drop the clause and thus save money on insurance. To overcome inefficiency, one needs to make sure that each contract is conditional on all counterparties in the system being insured, which can be thought of as equivalent to the insurance scheme set up above. However, it seems unreasonable that bilateral contracts requiring actions (insurance) from others than the two counterparties can be enforced.

3.3.3 A numerical example

For illustrational purposes I plug in the following numbers which seem to well describe the core of the financial system. By no means should these numbers be viewed as a calibration, since the model is too stylized to realistically describe the financial system. Think of the time frame of the model as being a few years, thus slightly longer than usual financial crises. I use $R = 1.01$, $R_{f,1} = 1$ to model that the assets of banks are not much more profitable than their liabilities, which in turn bear interest rates close to government bonds in normal times: it is the leverage that makes them profitable. I set $n = 50$ to model the core of the financial system and argue that while the number of large and entangled banks is relatively small worldwide, n is still large enough that there are severe externalities. The non-pledgable payoff is set to $X = 0.15$ reflecting that the stock market value of large banks is small compared to their balance sheets. The return to early liquidation is set to $L = 0.7$. The values related to the moral hazard are not observable, thus these are set to illustrate that the main point of the model works with reasonable values: loss from shirking is set to $d = 0.5$, private benefit in the initial period to $B_0 = 0.05$, in the interim period to $B_1 = 0.2$. The standard deviation of ϵ risks is set to $\sigma = 0.1$. I also assume that the financially sophisticated agents, the entrepreneurs, demand returns of $R_e = 1.25$. This value might seem high but it is a long-term return over more than a year, furthermore, the average annual return on common equity for investment banks was 16% before the crisis, so such high expected returns are not unreasonable.[17] Also, remember that this number also includes the costs of expertise, i.e. the wages of managers and bank employees. The probability of crisis in the baseline model is chosen to be $p = 0.05$, corresponding to crisis few times a century. With these values all the assumptions are satisfied.

The numerical values for the relevant variables are the following. The minimal equity share of a bank in the contagious equilibrium is $\underline{S}^h = 0.053$, which means the banks in the model work with about 20 leverage, which was not unusual for banks just before the crisis of 2008. The amount of debt with which a bank can survive the collapse of a neighbor is $\underline{S}^u = 0.088$, thus banks would have to increase their equity substantially to be stable, a decision which, given the pressure to produce profits to shareholders, they refuse to do.

Another major question is whether the failure of a single counterparty could prompt a run of wholesale creditors on a major financial institution. One example is Goldman: by the end of the year 2008 it already received payments of 12.9 billion from AIG after it was saved by the US government and the Federal Reserve Bank on September 16, 2008. The amount of its total stockholder equity on November 30, 2007 was 42.8 billion. Thus had AIG failed and had Goldman been unable to replace these contracts, it could have incurred substantial equity losses. While these numbers are not conclusive, it does illustrate that Goldman lenders fleeing as AIG was on the brink of failure might not have been that irrational after all.

3.4 Link with the reduced form model

This section shows that the same effects present in the reduced form model of Section 2 do exist in a more general microfounded framework. Here I show that in essence, the microfounded model and the reduced form model are very similar. The crucial step is to arrive at a VaR-like constraint for equity. From Lemma 2, the expression for the minimal amount of equity required to roll over debt can be split into two terms, a constant and another term depending on the default probability:

$$\underline{S}(R_{l,0}, \sigma_1) = \frac{B_1 - (R - R_{f,1}R_{l,0}) - X}{R_{f,1}R_{l,0}} + \frac{X}{R_{f,1}R_{l,0}} \cdot \underline{\pi}_2(R_{l,0}, \sigma_1)$$

The second term is a linear function of the default probability $\underline{\pi}_2$. Everything else given, the default probability at $t = 2$ is increasing in the amount of risk σ_1 held at $t = 1$. Even if this relationship is not linear, it is in essence very similar to the VaR constraint $S \geq \kappa\sigma_1$ of the reduced form model.

[17] "Capital spenders" in "A special report on international banking", The Economist, May 19th, 2007. The number is the average annual return on common equity of Bear Sterns, Goldman Sachs, Lehman Brothers, Merrill Lynch, Morgan Stanley and PaineWebber in the period 1985-2006.

3.5 Possible policy responses

3.5.1 Mandatory counterparty insurance

The most basic intervention is to introduce a mandatory counterparty insurance scheme to overcome the inefficiency highlighted in Proposition 2 if $p \in [p^s, p^i]$. There are two main mechanisms to do this: using a central insurance authority or a decentralized system of default swaps. The central insurance authority is basically equivalent to having a central clearinghouse that guarantees payments. Such a policy response has been proposed by several experts, such as Darrell Duffie[18], though he proposed these measures in order to eliminate counterparty credit risk and net out transactions. This paper shows that a similar response is needed in order to eliminate contagion through the loss of hedges.

The implementation of such a mandatory insurance scheme through a central counterparty can be done as follows. Set up a central insurance authority where a given amount of money has to be paid for each unit of bilateral exposure. This fund would then replenish the balance sheets of banks that lost their counterparties. The main drawback of such a setup would be that the central counterparty might not be able to price this insurance properly if there is disagreement about the exact probability p.

A decentralized mechanism is to let the market price default insurance in the form of default swaps. The regulation would then prescribe that for each unit risk insured with a counterparty (measured using e.g. the standard deviation of the risk), the bank has to buy counterparty insurance to cover the increase in equity it needs to roll over debt. In practice, for example for CDS's, this means the buyer of a credit default swap would have to insure against the default of the counterparty, the seller of the CDS, up to the notional amount on the CDS. Clearly this puts another potential layer of counterparty risk on top of the existing one. If banks want to game this regulation they could easily set up a firm, like a monoline insurer, who issues very cheap default insurance for all agents to cover the needs from regulation but then fails to pay in crisis. This is especially a problem given that for some intermediate values of p, banks do not find it individually optimal to participate in a counterparty insurance scheme (see Proposition 2). Thus in this decentralized setup there has to be strict enforcement to make sure default insurance, held against counterparty exposure, is highly diversified or that insurers do have very high equity reserves to cover even the worst potential losses.

3.5.2 Changing interest rates and recession insurance

Since the main mechanism of the breakdown of private counterparty insurance is due to the high costs of keeping reserves, one might think of other ways the regulator can de facto mandate holding slack equity. One way to achieve this is that in case a bank fails, the central bank eases the incentive constraints by lowering the return on the outside option: by setting a lower risk-free rate $R_{f,1}$ at $t = 1$. Note that this decrease has to be on rates of a maturity of a few years, which is problematic

[18] "Derivatives and Mass Financial Destruction: Complex financial markets can be useful is regulated properly" by Darrell Duffie, The Wall Street Journal, October 22, 2008

given that central banks only control short-term rates. If agents anticipate the central bank's move, they can now sign default swaps with each other to insure against counterparty risk, since they no more have to put aside equity to fulfill this promise but are provided equity in a state contingent way. Given that financial stability is not the central bank's sole responsibility, such a commitment might not be credible: central banks might want to increase interest rates to fight inflation or defend the currency, or already have interest rates that are close to zero. Given that such a policy response is not guaranteed, the counterparty insurance market may collapse again.

An approach that could be more credible, is one similar to the recession insurance for banks proposed Kashyap et al. (2008). The basic idea is that banks are mandated to take either of the following two actions. Either they pay insurance fees into a fund which pays them back in case of a recession in order improve their balance sheets, or hold higher equity reserves ex ante. This setup basically acts as a mandatory insurance scheme, thus making predictable payments to all banks which contrasts to the uncertain payments from cutting the risk-free rate. The shortcoming of this proposal for entangled systems is that it does not differentiate with respect to how entangled a given financial institution is, only by size of the balance sheet. Such heterogeneity in the number of OTC contracts of banks might call for a different policy: see Subsection 5 for a discussion of system with a large financial hub.

An interesting extension for future research would be to model the sales of illiquid real assets of the banks as they try to increase capital. Due to fixed investment size in this setup, I cannot model this or the effect of a central bank lending against illiquid assets, an important mechanism used by the Federal Reserve and the US Treasury in the crisis of 2008.

3.5.3 Nationalization

If there is a big outside player, e.g. the government, with financial expertise it could avert systemic crisis by nationalizing the financial institution which was hit by the shock. If it can manage the hedging contracts the bank signed with its counterparties then there will be no run on these neighboring banks. This solution also handles part of the moral hazard problem since shareholders are wiped out. Note however that realistically, managers and employees cannot be wiped out since they are needed to run the firm and manage its hedging contracts, leaving a substantial part of the moral hazard issue unresolved. The experience of the Federal Government with AIG in 2008-2009 shows that managing such an institution is far from easy and employees can still extract substantial amounts from the government even though shareholders are wiped out. Thus the ex-ante policy response of mandatory counterparty insurance is likely to dominate the ex-post response of nationalization.

3.5.4 Insurance provided by bondholders

Another potential insurance scheme would be one in which bondholders of the failed institution are forced to pay the neighbors of the failed bank out of the liquidation value L. Since bondholders anticipate the lower payout in crisis, they would increase interest rates, leading to slightly more capital being held in the system. However, this capital is only to offset expected losses from this insurance scheme, thus it is not as costly as holding enough capital in the system to replenish balance sheets. It is less costly, since bondholders store the equity and then provide it in a state-contingent fashion.

Even though this is a fairly unorthodox way to deal with the problem, it still could be viable. Problems with implementing such a policy are similar to that discussed in the case of renegotiation: the main benefit of a debt contract is its simplicity. Such a policy would make debt contracts extremely complicated state-contingent contracts. For the detailed argument see Subsection 3.2.7.

3.6 Debt covenants

In this subsection I argue that the inefficiency is still present even if we allow for debt covenants. I omit formal proofs, since these results are not central to the message of the paper. It is true though that debt covenants do mitigate inefficiency somewhat. To incorporate the idea of bond covenants in the model, I change the assumption about the financing choice in the following way. I assume banks can post a financing mix of equity S and debt $D = 1 - S$, and can credibly commit to it in the initial period. This means that investors can charge banks interest rates based on their future capital structure. Given that the investment and risk management decisions of the bank are complicated and hard to understand for outsiders, I do not allow for debt covenants prescribing ϵ hedging. However, allowing for such covenants would not change the analysis either: the inefficiency is due to the debt contract not being able to include effort choice as a covenant.

Probably the largest change is that long-term borrowing can no longer be ruled out as easily as in the case without debt covenants, since banks have a much stronger commitment device through committing to a given amount of equity. In case of long-term debt financing, the entrepreneur's problem changes to:

$$E_{t=0}\left[max(0, R_i + \epsilon_i - \epsilon_{i+1} - \delta_i + R_{f,1}G_i - R_L D_i + C_i)\right] + (1 - \pi_2)X - R_e S_i$$

s.t.

$$E_{t=0}\left[max(0, R_i + \epsilon_i - \epsilon_{i+1} - \delta_i + R_{f,1}G_i - R_L D_i + C_i)\right] + (1 - \pi_2)X - R_e S_i \geq 0$$

$$1 + G_i = S_i + D_i$$

$$E_{t=0}\left[max(0, R_i + \epsilon_i - \epsilon_{i+1} - \delta_i + R_{f,1}G_i - R_L D_i + C_i)\right] + (1 - \pi_2)X \geq B_0 + B_1$$

where the long-term lending rate R_L, is taken as given. The probability of survival at $t = 1$ is not included, since the endogenous probability of going bankrupt at $t = 1$ is zero, given that creditors do not have the option to withdraw funding in the interim period. The last constraint now ensures that the bank has the proper incentives to exert effort in both the initial and the interim periods, without the possible punishment of investors withdrawing financing at $t = 1$.

Why do banks choose short-term financing even if they can commit to a given capital structure, since it is exactly short-term financing that makes them vulnerable to withdrawal of debt funding and thus systemic crisis? Short-term debt is a good disciplining device to ensure high effort and appropriate hedging of risk. An important insight is that the higher the temptation to misbehave at $t = 0$, i.e. the higher B_0, the costlier is long-term financing compared to short-term debt financing. Thus if B_0 is high enough, banks do not choose long-term financing since it is too costly.

Once we ensure that banks choose short-term debt, the rest of the analysis does not change substantially. The only difference is that banks, when changing their capital structure do take into account that the interest rate charged to them changes as well. This leads only to second-order changes in the values of minimal equity, thus the basic qualitative insight of the model without debt covenants holds in this case too.

4 General equilibrium model and financial development

This section aims at endogenizing the network structure in a financial development framework. It shows that the sparse network that leads to inefficiency arises endogenously, since banks do not internalize the benefits of diversifying counterparty risk. It also highlights that the financial system is the most likely to be inefficient after an initial boom of financial development, once banks are completely connected but the network of bilateral claims is still sparse.

4.1 General equilibrium model setup

This section builds on the reduced form model of Section 2 with some changes and extensions. Each bank i can now choose the size of its real investment I_i. Instead of having unlimited capital and an outside option, each entrepreneur has limited capital of S. The accounting identity thus becomes $S + D_i = I_i$. Project i now has an expected a return of $I_i \cdot (R + \epsilon_i)$, thus the expected return and the risk grow linearly in the project size. The liquidation value also scales with investment size: $I_i \cdot L$. The risks are normally distributed independent random variables: $\epsilon_i \sim N(0, \sigma)$. If hit by the uninsurable idiosyncratic shock, bank i loses $d \cdot I_i$ of its equity capital. Again, $L > R - d$ is assumed to rule out self-insurance. The return on equity, which was exogenous in Section 2, is endogenous in this general equilibrium setup.

At an upfront cost of c (i.e. $\frac{c}{2}$ per bank), any two banks can establish a relationship, in which they learn about each others' risks, e.g. learn the distribution and can then sign any risk sharing contract. The number of counterparties of bank i is denoted by k_i. Note that the risk structure here is different from that in Section 2, since the risks of neighbors are not exactly offsetting, thus it does not automatically imply the sparse circular network of the baseline model. All other ingredients of the model are those described in the circular model in Section 2. The only assumption I make about how banks get connected is that if all of them have at least two connections, then all of them are linked together. The qualitative implications of the analysis would go through with random connections (in a random graph) too, so this assumption is only needed for simplicity.

4.2 General equilibrium model solution

4.2.1 Risk holding

Given that the setup is symmetric and there are no increasing returns to scale, I restrict my attention to symmetric equilibria where each bank has the same number of links: $k_i = k$ for all i. The exact number of k is determined in equilibrium. Given the number of counterparties, each bank diversifies its risk the most it can, since lower risk allows for scaling up the project and getting higher returns on equity. In equilibrium, if a bank chooses to diversify perfectly, it sells $\frac{1}{k+1}$ of the risk of its project to each of its counterparties and holds on to $\frac{1}{k+1}$ fraction of it.

Thus in equilibrium every bank holds $k+1$ uncorrelated risks with standard deviation of $I \cdot \frac{\sigma}{k+1}$ each. For example if bank 1 is connected to k banks $(2, 3, ..k+1)$, it holds the risk $\frac{I \cdot \epsilon_1}{k+1} + \frac{I \cdot \epsilon_2}{k+1} + .. + \frac{I \cdot \epsilon_{k+1}}{k+1}$. This amounts to a total risk with standard deviation of $I \cdot \frac{\sigma}{\sqrt{k+1}}$. Thus the VaR constraint for capital in normal times becomes:

$$S - \frac{k \cdot c}{2} \geq \kappa \cdot I \cdot \frac{\sigma}{\sqrt{k+1}}$$

where $S - \frac{k \cdot c}{2}$ is the capital left after paying for the fixed cost of establishing k counterparties.

If a bank loses one of its counterparties, e.g. bank $k+1$ in the previous example, the risk it holds after having lost a counterparty is $2\frac{I \cdot \epsilon_1}{k+1} + \frac{I \cdot \epsilon_2}{k+1} + .. + \frac{I \cdot \epsilon_k}{k+1}$, which has a standard deviation of $I \cdot \sigma \frac{\sqrt{k+3}}{k+1}$. Thus the bank needs the following amount of equity reserves to survive the loss of a counterparty:

$$\Delta S = \frac{\sqrt{k+3} - \sqrt{k+1}}{k+1} \cdot \kappa \sigma I$$

If all banks have k counterparties, so does the failed one, so the reserves in a counterparty insurance fund have to add up to:

$$A(k, I) = k \cdot \Delta S = \frac{k}{k+1}(\sqrt{k+3} - \sqrt{k+1}) \cdot \kappa \sigma I = \gamma I$$

where γ is introduced to simplify notation.

An alternative that has to be considered is that in expectation of the possible failure of a counterparty a bank decides to sell $\frac{1}{k}$ fraction of its risk to every counterparty and keep none of its risk on its own book. This way it has to hold $\kappa I \frac{\sigma}{\sqrt{k}}$ equity both if all its counterparties survive or if one of the counterparties fails since the amount of risk it holds does not change when the counterparty fails: in case of failure it just exchanges $\frac{1}{k}$ of its its failed counterparty's risk for its own risk. This equity is lower than that needed in the previous case if a counterparty fails, however all banks have to hold on to it even in normal times. In the Appendix, I show that this kind of insurance is definitely not efficient if $k \leq \frac{n}{2}$, i.e. for sparse networks that are our main interest, thus I can rule it out without changing the qualitative implications of the analysis.

4.2.2 Choice of the number of counterparties

The number of counterparties k is determined in equilibrium in the following way. First let us assume we are in the inefficient contagious equilibrium. Given initial capital of S and that each bank spent $\frac{k \cdot c}{2}$ of its capital to establish links with k counterparties, the banks can scale up to a size of:

$$I = \sqrt{k+1} \cdot \frac{S - \frac{k \cdot c}{2}}{\kappa\sigma}$$

The total payoff from investment of size I in a contagious system is:

$$IR - pI \cdot (R - L) - \left(I - \left(S - \frac{k \cdot c}{2}\right)\right) \cdot R_f$$

thus the banks are exactly indifferent between having k and $k+1$ links if and only if:

$$\sqrt{k+1} \cdot \frac{S - \frac{k \cdot c}{2}}{\kappa\sigma} \cdot (R - p \cdot (R - L) - R_f) + \left(S - \frac{k \cdot c}{2}\right) R_f =$$

$$\sqrt{k+2} \cdot \frac{S - \frac{(k+1) \cdot c}{2}}{\kappa\sigma} \cdot (R - p \cdot (R - L) - R_f) + \left(S - \frac{(k+1) \cdot c}{2}\right) R_f - \frac{c}{2}$$

Thus the banks will choose to have exactly k counterparties in a contagious system if both:

$$2\frac{R - p \cdot (R - L) - R_f}{\kappa\sigma} \frac{\sqrt{k+1} - \sqrt{k}}{\left(\sqrt{k+1} + (k-1) \cdot (\sqrt{k+1} - \sqrt{k})\right) \cdot \frac{R - p \cdot (R-L) - R_f}{\kappa\sigma} + R_f + 1} > \frac{c}{S}$$

and

$$\frac{c}{S} \geq 2\frac{R - p \cdot (R - L) - R_f}{\kappa\sigma} \frac{\sqrt{k+2} - \sqrt{k+1}}{\left(\sqrt{k+2} + k \cdot (\sqrt{k+2} - \sqrt{k+1})\right) \cdot \frac{R - p \cdot (R-L) - R_f}{\kappa\sigma} + R_f + 1}$$

are satisfied.

Figure 4 shows the relation between the cost of contracting and the number of counterparties. The curve is the indifference curve between having k and $k+1$ counterparties. In a comparative statics exercise, one can think of financial development as the decrease in the cost to equity of contracting, $\frac{c}{S}$. Once the cost is low enough, in the graph 3.3% of initial capital, but still relatively high, and banks start connecting to each other, the network is sparse for a wide range of costs: with the given parameter values there are less than 10 connections if the cost is above 1% of capital. This is similar to the well known result from diversification: most of the risk can be diversified with only a few uncorrelated investments. Only when the cost is substantially lower does the network start becoming dense. Ultimately, as the cost approaches zero, in the graph at 0.5%, the network becomes completely dense, thus in effect resulting in a complete market.

4.2.3 Inefficiency

The insurance equilibrium is not a subgame perfect Nash equilibrium if it is worth for one bank to opt out when all banks have k counterparties. Assume the size of the projects are I_{in} if insured and a single bank can scale up to I_{out} if it opts out of the counterparty insurance scheme. Notice though that when you opt out, your hedging contracts will be of the same size (since none of the other banks scale up in a one player deviation), thus the bank will bear all the residual risk it creates. The risk held by the deviating bank is:

$$\left((I_{out} - I_{in}) + \frac{I_{in}}{k+1}\right)\epsilon_1 + \frac{I_{in}}{k+1}\epsilon_2 + ... + \frac{I_{in}}{k+1}\epsilon_{k+1}$$

where w.l.g. bank 1 is the one deviating and its counterparties are banks $2, ..k+1$. Denote the increase in size by $\Delta I = I_{out} - I_{in}$. The standard deviation of the total risk held is now:

$$\sigma \cdot \sqrt{\frac{I_{in}^2}{k+1} + \Delta I^2 + 2\Delta I \frac{I_{in}}{k+1}}$$

Solving for ΔI using the the capital requirement, the positive root is:

$$\Delta I = \sqrt{\left(\frac{S - \frac{k \cdot c}{2}}{\kappa\sigma}\right)^2 - \frac{k \cdot I_{in}^2}{(k+1)^2}} - \frac{I_{in}}{k+1}$$

In the equilibrium where all banks insure and $A(k, I)$ capital is set aside in the counterparty insurance fund, the insured banks can scale up to the point that their capital requirement becomes binding:

$$S - \frac{A(k, I)}{n} - \frac{k \cdot c}{2} = \kappa \cdot \frac{I_{in}}{\sqrt{k+1}}\sigma$$

Using the expression for $A(k, I)$ we arrive at:

$$I_{in} = \frac{\sqrt{k+1}}{1 + \frac{k}{n}\left(\sqrt{\frac{k+3}{k+1}} - 1\right)} \cdot \frac{S - \frac{k \cdot c}{2}}{\kappa\sigma} = \beta \cdot \frac{S - \frac{k \cdot c}{2}}{\kappa\sigma}$$

Given the above definition of β we can rewrite the expression for ΔI, where we define α for the ease of notation:

$$\Delta I = \left(\sqrt{1 - \frac{k \cdot \beta^2}{(k+1)^2}} - \frac{\beta}{k+1}\right) \cdot \frac{S - \frac{k \cdot c}{2}}{\kappa\sigma} = \alpha \cdot \frac{S - \frac{k \cdot c}{2}}{\kappa\sigma}$$

A bank decides to opt out from counterparty insurance if and only if the expected payoff without insurance is higher than that with insurance. Note that if the bank is the only one to opt out it goes bankrupt if it or one of its counterparties is directly hit by the crisis, i.e. with probability $\frac{(k+1)\cdot p}{n}$:

$$(I_{in}+\Delta I)\cdot R-\frac{(k+1)\cdot p}{n}\cdot(I_{in}+\Delta I)\cdot(R-L)-\left(I_{in}+\Delta I-\left(S-\frac{k\cdot c}{2}\right)\right)\cdot R_f >$$

$$> I_{in}\cdot R-\frac{p}{n}\cdot I_{in}\cdot(R-L)-\left(I_{in}-\left(S-\frac{A(k,I)}{n}-\frac{k\cdot c}{2}\right)\right)\cdot R_f$$

Note that all above calculations assume $k \geq 2$, i.e. that the network is fully connected. For the special case of $k = 0, 1$, there is no externality and the social and private optimum coincide, furthermore there is no systemic crisis.

The insurance equilibrium is not sustainable in equilibrium if the probability of crisis is not high enough to prompt counterparties to insure:

$$p < p^{out} = \frac{\Delta I\cdot(R-R_f)+\frac{A(k,I)}{n}\cdot R_f}{\frac{k}{n}I_{in}\cdot(R-L)+\frac{k+1}{n}\Delta I\cdot(R-L)} = \frac{n\alpha\frac{R-R_f}{R-L}+\gamma\beta\frac{R_f}{R-L}}{k\beta+(k+1)\alpha}$$

On the other hand, the cutoff value for p in order for insurance to be socially optimal is lower. Note that if none of the banks insure but all hedge, then they can all scale up their investment to the maximum allowed by the VaR constraint:

$$I_{max} = \sqrt{k+1}\cdot\frac{S-\frac{k\cdot c}{2}}{\kappa\sigma}$$

Thus the social planner will choose insurance if and only if:

$$n\cdot I_{max}\cdot R-np\cdot I_{max}\cdot(R-L)-n\cdot\left(I_{max}-\left(S-\frac{k\cdot c}{2}\right)\right)\cdot R_f <$$

$$< n\cdot I_{in}\cdot R-p\cdot I_{in}\cdot(R-L)-n\cdot\left(I_{in}-\left(S-\frac{k\cdot c}{2}-\frac{A(k,I_{in})}{n}\right)\right)\cdot R_f$$

Note that the planner takes into account that in case of crisis in the uninsured network all n banks collapse, while in the insured one only a single one. Thus for $k \geq 2$, counterparty insurance is socially optimal if and only if:

$$p < p^s = \frac{n\cdot(\sqrt{k+1}-\beta)\frac{R-R_f}{R-L}+\gamma\beta\cdot\frac{R_f}{R-L}}{n\sqrt{k+1}-\beta}$$

Figure 5 shows the inefficiency region, which is the widest if the graph of bilateral contracts is sparse. It can even be considered sparse when each bank connect to a fifth or even a quarter of all banks in the system. Remember, that the network is endogenously sparse for an intermediate range of connection costs. From the perspective of financial development, the system is most likely to be inefficient and vulnerable to systemic crisis when financial development is in its early stage: when banks start connecting to each other.

One point highlighted by the graph is that the inefficiency result is probably most relevant to relatively infrequent crisis events, ones that happen a few times in a century (if one is willing to accept that the model incorporates a time frame of a few years). This crisis frequency does correspond to that of financial crises in developed economies. Even though these events are rare, it is socially optimal to hold reserves in an insurance fund all the time, since these funds used in an insurance scheme can then be used to avert a huge disaster.

Allowing for reinsurance does not change the qualitative findings, since the underlying reason for inefficiency is still present. Even if the k banks, whose counterparty went bankrupt, can reinsure at a cost of $\frac{c}{2}$ per bank, they still need to keep equity reserves to cover these costs. This is very similar to the case where banks whose counterparty has failed need additional equity reserves to be able to survive at $t = 1$. Once again banks have the incentive to free-ride on others. Thus in sparse networks, inefficiency is present even when allowing for reinsurance, although the inefficiency region could shift to lower levels of crisis probability.

Note that, as in the case of the circle, the failure of a single counterparty is enough to drive a bank into bankruptcy. This happens by the endogenous equity choice of banks. In a more general setting with heterogenous network structure some banks in a central position might hold enough equity to survive several counterparty defaults. A simple case of a financial hub (star network structure) is analyzed in Section 5.

4.3 Financial networks in 2008: dense or sparse?

The results of this section show that the crucial question when deciding whether the financial system is likely to be inefficient is the density of the hedging network. The theory shows that it is right after bilateral hedging appears in the financial system that the system is most likely to be sparse and thus inefficient. Was this the case in 2008? Unfortunately the data on bilateral hedging contracts of banks is not public. However, in early 2009, under public pressure because of the immense taxpayer funds it received, AIG did disclose the payments it made to counterparties on some bilateral hedging contracts. The payments to counterparties are a good proxy of the size bilateral exposure. Table 1 shows that these payments were heavily concentrated towards a few banks. The concentration of payments is clearly stronger than the concentration of financial institutions. The top three counterparties account for about half of the payments, hinting at a sparse network. One observation is not enough to conclude that the network is sparse but it shows that the regulator responsible for financial stability should collect information on the size of bilateral hedging exposure of banks to assess the stability of the financial system.

5 Centralized financial network

One might argue that the contracts and relationships in the OTC markets are not homogenous but the network structure is more centralized: e.g. there is a large financial hub (core) and the other smaller banks (periphery) do business with it instead of doing business with each other. Babus (2009) shows that such core-periphery network structures emerge endogenously in financial markets with strategic bilateral relationships. This section shows that the basic insight of inefficiency is present in such heterogenous networks as well: in general the core bank does not internalize the effect of its decisions on periphery banks thus it might not hold enough capital. It also highlights the fact that in a heterogenous system the banks' capital has to be regulated based on their own entangledness: more central banks should hold more capital.

5.1 Setup

To analyze the effect of node heterogeneity I consider a modification of the reduced form model discussed in Section 2: see Figure 6. The central hub (core) hedges risks with all other (periphery) banks. Periphery banks do not hedge with each other, thus we have a so-called star or core-periphery network structure. Periphery banks have unit projects as before, while the size of the real project of the central hub is $m \geq 1$ with a total return of $R_c = m \cdot \left(R - \frac{1}{m}\sum_{i=1}^{n} \epsilon_i\right)$ and the hub can contract on all ϵ risks and its liquidation value also scales linearly with its size. The real return of the $i = 1, 2..n$ peripheral banks is $R_i = R + \epsilon_i$ as before and bank i can only contract on ϵ_i. In the special case of $m = \sqrt{n}$, the variance of total risk of the real project to investment size is the same for the hub and the periphery banks. The distribution of the returns is the same as before. The given risk structure is by no means essential for the results but it simplifies the analysis substantially. I assume symmetry in bargaining over ϵ hedging contracts still holds, since not hedging has a similar equity cost on both the hub and the periphery bank in the equilibrium where all ϵ risks are hedged: thus all ϵ contracts have a price of zero as of $t = 0$ and $t = 1$.

Since the setup is asymmetric, I introduce two types of crises, which are mutually exclusive. The central hub fails with a probability of p_c, while one of the periphery banks fails with a probability of p_p, similarly to before. It is reasonable to assume $p_p \gg p_c$. The state of nature at $t = 1$ is "normal", thus no bank gets a shock to its equity, with probability $1 - p_p - p_c$. The worst possible idiosyncratic shock to equity of the periphery banks is d_p, while for the core it is $m \cdot d_c$. For both the periphery bank and the hub assume that it is socially optimal to liquidate a bank hit by the idiosyncratic shock: $L > R - d_p$ and $L > R - d_c$.[19]

5.2 Underinsurance and inefficiency

The hub needs equity of $\kappa\sigma$ to survive at $t = 1$ if a periphery bank collapses. Since the hub is connected to all periphery banks, it perceives the probability of needing the equity to be p_p (contrast this to the case of a circle, where it was $\frac{2p}{n}$). The hub finds it individually optimal to keep enough equity to

[19]This assumption is made for simplicity, it rules out the case that banks want to hold enough equity to self-insure against the idiosyncratic shock in equity.

survive a periphery default if and only if:

$$mR - p_c \cdot m \cdot (R - L) - (m - \kappa\sigma) \cdot R_f - \kappa\sigma \cdot R_e \geq mR - (p_c + p_p) \cdot m \cdot (R - L) - mR_f$$

thus the minimum level of p_p to prompt the hub to insure given its individual incentives is:

$$p_p \geq p_p^i = \frac{\kappa\sigma}{m} \cdot \frac{R_e - R_f}{R - L}$$

Along similar lines, one gets that periphery banks only hold equity against the failure of the central hub if:

$$p_c \geq p_c^i = \kappa\sigma \cdot \frac{R_e - R_f}{R - L}$$

In this latter case, all periphery banks hold equity of $\kappa\sigma$ each, thus they do not even have to sign the ϵ hedging contract with the hub: they choose autarchy. Thus for a range of parameter values, the central hub holds equity to insure against the collapse of a periphery bank but the periphery banks do not hold equity to survive the collapse of the central hub. This is likely to be the case if p_c is substantially smaller than p_p and/or m is large. It just remains to show that periphery banks, do not insure if the central hub does not insure against the failure of periphery banks, since the above derivation implicitly assumes this. Periphery banks would insure if and only if:

$$R - \frac{p_p}{n} \cdot (R - L) - R_f - \kappa\sigma(R_e - R_f) \geq R - (p_c + p_p) \cdot (R - L) - R_f$$

which yields:

$$\frac{n-1}{n} p_p + p_c \geq \kappa\sigma \cdot \frac{R_e - R_f}{R - L}$$

Thus it can never be the case that the central hub does not insure while periphery banks do since $\frac{n-1}{n} < m$.

5.3 Optimal policy

On the other hand, the socially optimal outcome would be for the core to insure against periphery failure already at a lower probability of periphery collapse, if and only if:

$$(m+n)R - p_c \cdot (m+n) \cdot (R-L) - (m+n)R_f - \kappa\sigma \cdot (R_e - R_f) \geq (m+n)R - (p_c + p_p) \cdot (m+n)R - (m+n)R_f$$

yielding:

$$p_p \geq p_p^s = \frac{\kappa\sigma}{m+n} \cdot \frac{R_e - R_f}{R - L}$$

which is clearly lower than p_p^i, especially if e.g. $m = \sqrt{n}$. Thus there is inefficiency due to the hub not taking into account that if it fails due to a periphery bank going bankrupt, all periphery banks go bankrupt. However, unless p_p is very low, this is unlikely to be the problem.

However, there is another more likely inefficiency. Given the centrality of the hub, it is in some cases socially optimal to regulate the hub and make sure it has enough equity to even survive the idiosyncratic shock itself. Remember that given the assumption of $L > R - d_c$, it is never individually optimal for the central hub to hold enough equity to survive the idiosyncratic shock to equity.[20] Forcing the hub to keep equity of d_c is socially optimal if and only if:

$$(m+n)R - p_c m d_c - p_p(R-L) - (m(1-d_c)+n)R_f - m d_c R_e >$$

$$> (m+n)R - p_c m(R-L) - p_p(R-L) - p_c n(R-L) - ((m-\kappa\sigma)+n)R_f - \kappa\sigma R_e$$

where on the right hand side $-p_c n(R-L)$ is the extra loss not internalized by the hub that all other banks fail if it fails. I also assumed in the above calculation that the core has enough capital to survive the collapse of a single periphery bank, this assumption does not influence the result though. The calculation yields:

$$p_c > p_c^{s,ih} = \frac{d_c - \frac{1}{m}\kappa\sigma}{\frac{n}{m}(R-L) + (R-L-d_c)}(R_e - R_f)$$

This means that the regulator might find it socially optimal to force the hub to maintain very large amount of equity to protect the rest of the financial system from potential mistakes or fraud in the bank acting as a hub.

To give a numerical example of which kind of inefficiency is likely to occur, take the following parameter values: $R_r = 1.01$, $R_e = 1.15$, $R_f = 1$, $n = 50$, $m = \sqrt{n}$, $L = 0.7$, $d_c = d_p = 0.4$, $\sigma = 0.2$, $\kappa = 2$. With these values, the hub finds it individually optimal to insure against the failure of a periphery bank if $p_p > p_p^i = 2.7\%$, while the social optimum would be to insure already if $p_p > p_p^s = 0.33\%$. Thus if there are many periphery banks and a periphery failure is likely then this form of inefficiency is unlikely. On the other hand, the regulator should force the hub to hold 40% equity (from $d_c = 0.4$) if $p_c > p_c^{s,ih} = 2.7\%$ even though the hub would individually never be willing to hold such a huge equity buffer. However, if the regulator fails to insure the hub has a large equity buffer, then the periphery banks are not willing to insure against the failure of the hub (i.e. not contract with it) unless $p_c > p_c^i = 0.19$. Thus only if the probability of the hub failure is very high, do periphery banks refuse to connect to the hub. This shows that the intuition of Warren Buffett on banks trying to minimize their links in a contagious system does not apply in general (see Subsection 2.3 for the quote).

The analysis in this Section highlights the fact that an asymmetric network structure needs very different policy response than the homogenous network. It is in general not optimal here for all banks to hold equity for the case that the central hub collapses, rather the central hub has to hold a huge equity reserve based on it being very entangled. Such policy has to depend on the degree of entangledness, not simply balance sheet size. This insight is clear from the setup, since in the model, the balance sheet size of the central hub m can be varied freely depending on what technology it has

[20] See Subsection 2.2 for the proof

to transform risk. Since regulating the hub benefits the periphery, the regulator might consider taxing them and giving a subsidy to the hub to counterbalance the cost that it has to hold high equity. The question of whether or not the periphery should be taxed cannot be answered in this framework since we assumed equal bargaining power of the periphery and core banks. In general, one might expect the core bank can extract some of the surplus from the periphery banks, thus the optimal policy regarding redistribution depends on the bargaining power of the periphery banks.

I restrict my attention to a star network structure because general heterogenous networks are difficult to analyze, since the banks can make very general choices: e.g. banks with many connections and large projects might opt to keep equity to bail out smaller, less connected, banks in order to prevent domino effects among these institutions. This is the case e.g. for a large hub bank linked to all periphery banks which are linked together on a circle. However, once again free riding problems can arise if there are several well connected banks. Another difficulty is that connected banks expect domino effects to happen and while it might be worthwhile to keep equity for the failure of a few counterparties, this might not be worthwhile any more once the bank expects widespread defaults among financial institutions. While these questions are very important and a promising avenue for future research, they are beyond the scope of this paper.

5.4 Financial hubs and the crisis of 2008

In some sense this setup might even be considered more relevant to the cases of Bear Stearns, Lehman Brothers, and AIG. In this interpretation the ex ante policy regarding the regulation of these connected financial giants failed in not imposing higher capital requirements on them. Had these central institutions held high equity buffers in 2008, the Federal Reserve would not have had to pump huge amounts of equity into the financial system: a policy which was ex ante suboptimal, but ex post optimal in order to avoid the collapse of all periphery banks.

If the financial hub can provide equity free, i.e. at $R_e = R_f$, then the hub holds ample equity and efficiency is restored. One can think of this as the Federal Reserve acting as a clearinghouse: it can provide unlimited equity since it can print money. This is basically the role the Federal Reserve assumed in 2008 by e.g. in effect nationalizing AIG and guaranteeing its contracts. Note however that it is likely that the Federal Reserve does not have the proper expertise to manage such a financial business, since it is hard to price the hedging contracts: in this case ex ante regulation should assure that private banks achieve the efficient outcome.

6 Concluding remarks

This paper develops a model of an entangled financial system where banks use bilateral contracts to hedge their risks. In such a system, the failure of a single bank can cause the whole system to collapse, even though there are no direct credit exposures between the banks. The mechanism of contagion proposed in this model is novel. The default of a neighboring bank makes a bank risky, decreasing its probability of survival. Since non-pledgable continuation value is used by banks as collateral in their incentive problems, this leads short-term creditors to stop lending to the bank, forcing it into bankruptcy. If all banks are connected, this mechanism leads to a complete collapse of the financial system as creditors run on all banks simultaneously once there is news about insufficient equity at a single bank.

I show that banks have an incentive to use bilateral contracts to hedge risks and thereby expose themselves to counterparty risk. The banks are unwilling to participate in an insurance scheme that would stop crisis from spreading. The main insight is that there is a market failure: the externalities of bankruptcy inflicted on others through derivative contracts is not internalized by the banks. Banks engage in free-riding: every bank would prefer that the others pay the costs of stability, not contributing itself.

The paper has important policy implications. Under reasonable conditions, making counterparty insurance mandatory would be welfare improving: it would prevent the financial system from collapsing in the low probability event of a crisis. Such an insurance scheme could for example be implemented by setting up a strongly regulated central counterparty that guarantees trades even if a large institution fails. The paper also calls for further research on regulating a general financial system with heterogenous structure of bilateral hedging. The model is a first step towards modeling the modern financial system which has been described as "entangled", in the sense that risk-hedging is based on a network of bilateral contracts.

Appendix: Proofs omitted in text

Proof of Lemma 1

Proof. Assume $G_i > 0$. I show that bank i always finds it weakly preferable to decrease leverage by decreasing the amount of borrowed D_i and pushing down liquid reserves G_i to zero. Note that the only way final payoff conditional on survival depends on G_i and D_i is through the term $R_{f,1}G_i - R_{l,1}R_{l,0}D_i$. Using the accounting identity to drop $D_i = 1 + G_i - S_i$ this becomes

$$(R_{f,1} - R_{l,1}R_{l,0}) \cdot G_i - R_{l,1}R_{l,0} \cdot (1 - S_i)$$

Given that the probability of default of any institution is non-negative $R_{l,1}R_{l,0} \geq R_{f,1}$ since in case of default not the whole amount of debt will be recovered so the interest rate on the loan must be higher than the risk-free rate. Thus choosing $G_i = 0$ is weakly preferable (strictly if and only if bankruptcy occurs in equilibrium) for payoff in case of survival.

Notice that the probability of survival also increases if the distribution of final payoff is shifted toward higher payoffs, thus choosing $G_i = 0$ is optimal given the objective function and the accounting identity. The incentive constraint and the participation constraint also depend on the final payoff in the same way as the objective function, thus choosing $G_i = 0$ is optimal for banks. □

Proof of Lemma 2

Continued from the text.

Proof. Taking the derivatives of the three equations determining the equilibrium, one can solve for the partial derivative of interest:

$$\frac{\partial \underline{S}}{\partial \sigma_1} = \frac{X}{R_{l,0}R_{f,1}} \cdot \frac{\Phi(x) \cdot \overbrace{(R - R_{l,0}\underline{R}_{l,1}(1 - \underline{S}))}^{>0} + \phi(x)\sigma}{\sigma^2 \frac{\Phi(x)}{\phi(x)} + \sigma X} > 0$$

where $(R - R_{l,0}\underline{R}_{l,1}(1 - \underline{S})) > 0$ since the pledgable income has to be larger than the total repayment on debt. This means that both the denominator and the nominator are positive, thus the partial derivative is positive. □

Proof of Lemma 5

Proof. If the bank takes long-term debt without any covenants, it can simply keep no equity and exert low effort in both periods to receive private benefits of $B_0 + B_1$. This is a lower bound on the payoff it can get from shirking. On the other hand one can derive an upper bound on payoffs if it behaves. Since equity is costly, $S = 0$ gives an upper bound on equity. The lower bound on interest rates from

$t = 0$ to $t = 2$ is the riskless rate on government bonds $R_{f,1}$. Thus the upper bound on payoffs with behavior and long-term debt is $(R - R_{f,1}) + X$, which by Assumption 3 is lower than $B_0 + B_1$. Thus the bank has no incentive to exert effort with long-term debt.

We also have to check wether short-term debt is feasible. Choosing short-term debt the bank can still choose to hold $S = 0$, misbehave and then go bankrupt at $t = 0$. The payoff from this behavior is B_0. Assuming the complete collapse of the network in crisis, which basically gives a lower bound on profits with short-term debt and exerting effort, the expected profit is:

$$E[P] = (1-p)[R + X - R_{f,1}R_{l,0}\underline{D}^h(R_{l,0})] - R_e\underline{S}^h(R_{l,0}) = (1-p)B_1 - R_e\left(\frac{B_1 - R - X}{R_{f,1}R_{l,0}^h} + 1\right)$$

Substitute $\underline{D}^h(R_{l,0})$ and notice that since at least some of the debt is recovered even in crisis: $R_{l,0}^h < \frac{1}{1-p}$. The condition for exerting effort at $t = 0$ when holding short-term debt is:

$$(1-p)B_1 - R_e\left((1-p)\frac{B_1 - R - X}{R_{f,1}} + 1\right) > B_0$$

which is satisfied by Assumption 5. □

Proof of Lemma 6

Proof. The potential gain from not hedging ϵ risks comes from two sources. First, equityholders can shift risk to debtholders. Second, by not hedging they could escape the contagion in systemic crisis. I show that there are no gains through the above two channels. On the other hand, not hedging can potentially decrease the expected value of non-pledgable payoff through increasing the probability of bankruptcy at $t = 2$, it is not be chosen in equilibrium.

Since ϵ contracts are observable at $t = 1$, if the bank does not hedge in order to shift risk to debtholders, it will face higher borrowing costs at $t = 1$ when the short-term debt is renewed, which exactly offsets the gains from risk-shifting. For the same reason, ϵ contracts are chosen to be long-run contracts to avoid penalty interest rates by lenders at $t = 1$.

The second possible gain from not hedging ϵ risks is to escape contagion in crisis. When the bank's neighbor defaults on its bilateral ϵ hedging contract, there are no direct losses, contagion spreads through losing the ϵ contract itself and becoming risky. However, even when banks hold ϵ hedging contracts, they could just as well hold the same amount of equity they would have as a risky autarkic bank without ϵ hedges and be resilient to contagion. If they chose to hold lower equity in case they are hedged, it is exactly because that increases their expected payoff. Thus there cannot be any gain from not hedging ϵ risks. □

Proof of Proposition 1

Proof. If no one holds slack equity, no single bank has the private incentive to deviate and hold equity that is just enough to survive the collapse of the rest of the system if:

$$(1-p)[R - R_{f,1}R^h_{l,0}\underline{D}^h(R^h_{l,0}) + X] - R_e\underline{S}^h(R^h_{l,0}) \geq$$

$$\geq (1-p)[R - R_{f,1}R^h_{l,0}\underline{D}^a(R^h_{l,0}) + X] + \frac{n-1}{n}p[R - R_{f,1}R^h_{l,0}\underline{D}^a(R^h_{l,0}) + (1-\tilde{\pi}_2)X] - R_e\underline{S}^a(R^h_{l,0})$$

where

$$\tilde{\pi}_2 = \Phi\left(\frac{\tilde{R}_{l,1}R^h_{l,0}\underline{D}^a(R^h_{l,0}) - R}{\sqrt{2}\sigma}\right)$$

and $\tilde{R}_{l,1}$ is set s.t. the investors break even. The left hand side of the inequality is the expected payoff not holding reserves, and the right hand side is that when the bank holds enough equity $\underline{S}^a(R^h_{l,0})$ to survive a crisis. Note that in the equilibrium creditors expect all banks to fail in crisis and charge interest rates accordingly. They charge the same initial interest rate to every bank, even to the one that chooses to deviate and hold higher equity. This is because the deviating bank cannot commit to holding higher equity either. A bank with equity $\underline{S}^a(R^h_{l,0})$ only goes bankrupt if it is directly hit by the adverse idiosyncratic shock at $t = 1$ or in case it falls short of paying back debts at $t = 2$ in a systemic crisis. Note that in the latter case the bank is risky, so the interest rate on its loan will jump to $\tilde{R}_{l,1} > R_{f,1}$. Rearranging and using that the incentive constraint is binding yields:

$$R_e[\underline{S}^a(R^h_{l,0}) - \underline{S}^h(R^h_{l,0})] + (1-p)\underbrace{R_{f,1}[R^h_{l,0}\underline{D}^a(R^h_{l,0}) - R^h_{l,0}\underline{D}^h(R^h_{l,0})]}_{-\tilde{\pi}_2 X} \geq \frac{n-1}{n}p\underbrace{[R - R_{f,1}\underline{R}_{l,0}\underline{D}^a(R^h_{l,0}) + (1-\tilde{\pi}_2)X]}_{B_1}$$

rearranging we arrive at the implicit expression for p^a:

$$p^a = \frac{R_e[\underline{S}^a(R^h_{l,0}) - \underline{S}^h(R^h_{l,0})] - \tilde{\pi}_2 X}{\frac{n-1}{n}B_1 - \tilde{\pi}_2 X}$$

However, the bank could choose to hold even more equity than that needed to simply roll over its debt if it is worth to increase the probability of survival. The marginal increase in survival probability at $t = 2$ in case more equity is held:

$$\frac{\partial \pi_2}{\partial S} = -\frac{\frac{1}{\sigma_1}\phi(x)R_{f,1}R_{l,0}}{1-\Phi(x)}$$

where $x = \frac{R - R_{l,1}R^h_{l,0}(1-S)}{\sigma_1}$. The expected payoff given equity S is:

$$E[P] = (1-p)\cdot[R - R_{f,1}R^h_{l,0}\cdot(1-S) + X] + \frac{n-1}{n}p\cdot[R - R_{f,1}R^h_{l,0}\cdot(1-S) + (1-\pi_2(S))\cdot X] - R_e S$$

where the second term is the payoff in case the whole financial system collapses except for the the one holding surplus equity. The change in $E[P]$ if the bank increases S over the minimum level needed for rollover:

$$\frac{\partial E[P]}{\partial S} = \left(1 - \frac{p}{n}\right) \cdot R_{f,1}R_{l,0}^h - \frac{n-1}{n} pX \frac{\partial \pi_2}{\partial S} - R_e$$

note that $R_{l,0}^h$ does not change if the bank chooses higher equity, since at $t = 0$ it cannot commit to higher equity. Given the costs of holding more equity, the bank will choose not to overinsure if and only if:

$$\frac{n-1}{n} pX \left(-\frac{\partial \pi_2}{\partial S}\right) \leq \frac{p}{n} R_{f,1}R_{l,0}^h + (R_e - R_{f,1}R_{l,0}^h)$$

Substituting $\frac{\partial \pi_2}{\partial S}$, it is sufficient to show that:

$$pX \frac{1}{\sigma_1} \frac{\phi(x)}{1 - \Phi(x)} \leq \frac{R_e}{R_{f,1}} \left(1 - \frac{p}{n} L\right) - 1$$

since $\frac{\phi(x)}{1-\Phi(x)} \leq \sqrt{\frac{2}{\pi}}$, and $\sigma_1 = \sqrt{2}\sigma$, this holds by Assumption 6. Thus the bank does not choose to hold reserves beyond that needed to roll over debt.

We now pin down the interest rate on the bond from $t = 0$ to 1. In the short run investors demand expected return of 1. Given that all banks are liquidated in crisis the indifference condition is:

$$1 \cdot \underline{D}^h(R_{l,0}) = (1 - p) \cdot R_{l,0}\underline{D}^h(R_{l,0}) + p \cdot L$$

Substituting $\underline{D}^h(R_{l,0}) = \frac{R+X-B_1}{R_{f,1}R_{l,0}}$ the above expression yields the interest rate stated in the lemma.

Now we turn to showing that the participation constraint of the entrepreneur $E[P] > 0$ is satisfied if all banks hold this minimal equity using ϵ hedging and short-term debt is the following. Given the complete collapse of the network in crisis, which basically gives a lower bound on profits, the expected profit is:

$$E[P] \geq (1-p)[R + X - R_{f,1}R_{l,0}^h\underline{D}^h(R_{l,0}^h)] - R_e\underline{S}^h(R_{l,0}^h) = (1-p)B_1 - R_e \left(\frac{B_1 - R - X}{R_{f,1}R_{l,0}^h} + 1\right)$$

Since at least some of the debt is recovered even in crisis: $R_{l,0}^h < \frac{1}{1-p}$, the condition for participation simplifies to

$$(1-p)B_1 - R_e \left(1 - (1-p)\frac{R + X - B_1}{R_{f,1}}\right) > 0$$

which is satisfied by Assumption 5, given that $B_0 \geq 0$. Thus entrepreneurs choose to participate in such a contagious system. □

Proof of Lemma 7

Proof. The main problem with any insurance scheme is that it undermines effort. Now assume there is an insurance scheme in place that guarantees bank i to continue at least until $t = 2$ if the signal about its idiosyncratic shock was bad, i.e. $s_i = d$. Note that the insurance scheme cannot be made contingent on effort, only on this signal, which may either be due to an adverse shock or to low effort at $t = 0$. Thus if a bank chooses low effort at $t = 0$, it is guaranteed to continue to $t = 2$, thus it can choose low effort again. Thus by low effort and an initial choice of equity $S_i = 0$ it can achieve payoff of $B_0 + B_1$. On the other hand if it chooses high effort, the upper bound on its profits is: $R - R_{f,1} + X$ where we assumed it could borrow all fund for the investment at the riskless late, and that the investment succeed with probability one. Clearly, by Assumption 3, it is a profitable deviation to shirk, thus insuring against $s_i = d$ is not feasible. $\square$

Proof of Proposition 2

Proof. The insurance pool needs reserves adding up to $2[\underline{S}^u(R_{l,0}) - \underline{S}^h(R_{l,0})]$, given the interest rate $R_{l,0}$, to be able to stabilize the two neighbors of the failing bank by granting them additional equity. For now I simply assume that the banks getting the capital injection only get just enough to be able to roll over debt: I show that this is indeed the case at the end of this proof. Assume the counterparty insurance is in place at a cost of k per bank. The private incentive of a single bank to deviate from counterparty insurance (opt-out), thus become uninsured in a system where the rest of the network is insured and thus stable, is:

$$\left(1 - 3\frac{p}{n}\right)[R - R_{f,1}\hat{R}_{l,0}\underline{D}^h(\hat{R}_{l,0}) + X] - R_e\underline{S}^h(\hat{R}_{l,0}) \geq$$

$$\geq \left(1 - \frac{p}{n}\right)\left[R - R_{f,1}\hat{R}_{l,0}\underline{D}^h(\hat{R}_{l,0}) + X\right] - R_e\underline{S}^h(\hat{R}_{l,0}) + R_{f,1} \cdot \underbrace{\frac{2p}{n}\left(\underline{S}^u(\hat{R}_{l,0}) - \underline{S}^h(\hat{R}_{l,0})\right)}_{\text{expected insurance receivable}} - 2\frac{p}{n}\hat{\pi}_2 X - k$$

where the probability of failure at $t = 2$ in case a bank's neighbor fails and it gets capital infusion from the insurance is:

$$\hat{\pi}_2 = \Phi\left(\frac{\hat{R}_{l,1}\hat{R}_{l,0}\underline{D}^u - R}{\sigma}\right)$$

and $\hat{R}_{l,0}$ is the endogenous short-term rate if all banks choose to insure and $\hat{R}_{l,1}$ is the interest rate faced by the two banks that get capital injection in the crisis state. If the bank is uninsured but all the others are insured, it goes bankrupt if either it or its two neighbors are hit by the adverse idiosyncratic shock in crisis, thus with probability $\frac{3p}{n}$. By Lemma 3, if financing is withdrawn at $t = 1$, then the equity holder does not get anything. In case of choosing to be insured, the term $-2\frac{p}{n}\hat{\pi}_2 X$ has to be added, since even if a bank's neighbor failed and it has $\underline{S}^u(\hat{R}_{l,0})$, i.e. can survive at $t = 1$, you may still fail at $t = 2$ and lose the non-pledgable payoff. Since the investors cannot observe the deviation from insurance when they lend to the bank they cannot charge higher interest rates to the deviating bank. Note that having extra equity funds in the insurance vehicle does not allow the banks to borrow more, since in crisis these reserves will be all transferred to the banks neighboring the one in trouble and all other banks have to be able to roll over their debt based on the equity on their own balance sheet.

The cost k of insurance per bank has to cover the opportunity cost and the expected payout of the insurance fund:

$$k = \frac{1}{n}\left[2\left(\underline{S}^u(\hat{R}_{l,0}) - \underline{S}^h(\hat{R}_{l,0})\right)(R_e - R_{f,1}) + R_{f,1} \cdot \underbrace{2p\left(\underline{S}^u(\hat{R}_{l,0}) - \underline{S}^h(\hat{R}_{l,0})\right)}_{\text{expected payout of fund}}\right]$$

We arrive at an implicit equation for p^i:

$$p^i : \qquad p = \frac{\left(\underline{S}^u(\hat{R}_{l,0}) - \underline{S}^h(\hat{R}_{l,0})\right)(R_e - R_{f,1})}{R - R_{f,1}\hat{R}_{l,0}\underline{D}^h(\hat{R}_{l,0}) + X - \hat{\pi}_2 X} = \frac{\hat{\pi}_2 X(R_e - R_{f,1})}{R_{f,1}\hat{R}_{l,0}(B_1 - \hat{\pi}_2 X)}$$

The way to calculate social welfare is the following. The expected payoff of debtholders is the same irrespective of the equilibrium, since they set interest rate to ensure they get the risk-free return in expectation. Thus we only have to look at the expected payoff of the entrepreneurs. The probability of failure π now has to be split into the probability of failure at $t = 1$ of π_1, and that at $t = 2$ of π_2, thus $\pi = \pi_1 + \pi_2$. The expected payoff of a single entrepreneur is, in general:

$$W = (1 - \pi_1)[R + X - R_{f,1}R_{l,0}D] - \pi_2 X - R_e S$$

where we used that creditors receive an expected return of $R_{f,1}$ at $t = 1$ on their investment of $R_{l,0}D$ but if the bank fails with probability π_2, it loses the non-pledgable payoff. Now we use that $R_{l,0}$ is set such that creditors break even in expectation at $t = 0$:

$$D = (1 - \pi_1)R_{l,0}D + \pi_1 L$$

Substituting $D = 1 - S$ on the left hand side and rearranging this yields:

$$(1 - \pi_1)R_{f,1}R_{l,0}D = R_{f,1} - R_{f,1}S - \pi_1 R_{f,1}L$$

which we now substitute into the welfare W:

$$W = (1 - \pi_1)(R + X) - \pi_2 X + \pi_1 R_{f,1}L - R_{f,1} - (R_e - R_{f,1})S$$

In the insured equilibrium the probability of going bankrupt at $t = 1$ is that of being hit directly by the idiosyncratic shock, $\pi_1 = \frac{p}{n}$. The probability of going bankrupt at $t = 2$ is that of being the neighbor of the failed bank, times the conditional probability of going bankrupt with equity $\underline{S}^u(\hat{R}_{l,0})$ and a missing ϵ hedge, thus $\pi_2 = 2\frac{p}{n}\hat{\pi}_2$. The social benefits outweigh the social costs of the counterparty insurance for the system as a whole, i.e. for all entrepreneurs, if:

$$n(1-p)[R+X]+npR_{f,1}L \leq (n-p)[R+X]+pR_{f,1}L-2p\hat{\pi}X-2\left(\underline{S}^u(\hat{R}_{l,0})-\underline{S}^h(\hat{R}_{l,0})\right)(R_e-R_{f,1})$$

where we simplified by $n\underline{S}^h(\hat{R}_{l,0})\cdot(R_e-R_{f,1})$. The left hand side is the total expected payoff to all market participants in case of no insurance, given that in case of a crisis there is a complete liquidation of all real projects. The right hand side is the total expected payoff in case the insurance scheme is in place, regardless of whether that constitutes an equilibrium. Thus counterparty insurance is socially optimal if $p>p^s$, where the implicit equation for p^s is:

$$p^s: \qquad p=\frac{2}{n-1}\frac{\left(\underline{S}^u(\hat{R}_{l,0})-\underline{S}^h(\hat{R}_{l,0})\right)(R_e-R_{f,1})}{R+X-R_{f,1}L-\frac{2}{n-1}\hat{\pi}_2X}$$

where substituting $\underline{S}^u(\hat{R}_{l,0})-\underline{S}^h(\hat{R}_{l,0})=\frac{\hat{\pi}_2X}{R_{f,1}\hat{R}_{l,0}}$ yields the expression sated in the lemma.

To prove that there is no counterparty insurance if $p\in(p^s,p^i)$ even though it is socially optimal, we have to show the impossibility of a mixed equilibrium where some insure and others do not. The key insight is that the amount of reserves needed to stop the spread of the crisis is the same irrespective of how many banks contribute. Thus if only every second bank contributes, then they perceive the probability of being effected as double compared to the system where everyone insures, doubling the expected gains from insurance. However, since in this mixed equilibrium only every other bank contributes, the cost of insurance doubles too. Another subtle point is that if not all banks insure, then in crisis a given portion of the network collapses. Since investors break even on debt and anticipate the equilibrium, the initial interest rate $R_{l,0}$ is higher in a mixed equilibrium than in the full insurance equilibrium. This means more equity has to be held since the probability of default, ceteris paribus, increases in $R_{l,0}$. Thus the mixed equilibrium is only an equilibrium if the full insurance is an equilibrium. The same argument holds for any mixed equilibrium, thus we can rule them out.

Up to now I simply assumed that the banks getting the capital injection from the insurance fund receive enough only to just about roll over their debt at $t=1$. Here I explore whether the banks would possibly like to transfer more equity to the two banks. The expected payoff for an individual bank under the insurance scheme is:

$$E[P]=\left(1-\frac{p}{n}\right)\left[R-R_{f,1}\hat{R}_{l,0}\underline{D}^h(\hat{R}_{l,0})+X\right]-R_e\underline{S}^h(\hat{R}_{l,0})+R_{f,1}\cdot\frac{2}{n}\left(\underline{S}^u(\hat{R}_{l,0})-\underline{S}^h(\hat{R}_{l,0})\right)-2\frac{p}{n}\hat{\pi}_2X$$

Increasing the amount of equity in the insurance fund by increasing $S^u(\hat{R}_{l,0})$, has the following marginal effect on expected payoff:

$$\frac{\partial E[P]}{\partial S^u}=-\frac{2}{n}\cdot(R_e-R_{f,1})-\frac{2p}{n}X\frac{\partial\hat{\pi}_2}{\partial S^u}$$

where $\sigma_1=\sigma$, since the neighbors of the failed bank lose only one of their two hedges. Following the similar derivation in Proposition 1, we arrive at the following sufficient (but not necessary) condition for the banks not to choose to increase the insurance fund:

$$pX\frac{1}{\sigma}\sqrt{\frac{2}{\pi}} \leq \left(\frac{R_e}{R_{f,1}} - 1\right) \cdot \left(1 - \frac{p}{n}L\right) - 1$$

which is fulfilled by Assumption 6. □

Proof of statement in 4.2.1

Proof. In the equilibrium where banks sell off all their own risk (I call this over-diversification), they keep equity of $S = \kappa I_{od}\frac{\sigma}{\sqrt{k}}$, thus given that their equity is fixed, they can scale up their investment to $I_{od} = \frac{\sqrt{k}S}{\kappa\sigma}$. In the insurance equilibrium, the equity needed by the banks after contributing to the central insurance pool is:

$$S - \frac{1}{n}\frac{k}{k+1}(\sqrt{k+3} - \sqrt{k+1})\kappa\sigma I_{od} = \kappa \cdot \frac{I_{in}\sigma}{\sqrt{k+1}}$$

thus the investment projects can expand to the size of:

$$I_{in} = \frac{S}{\kappa\sigma}\frac{1}{\frac{n-k}{n}\frac{1}{\sqrt{k+1}} + \frac{k}{k+1}\frac{\sqrt{k+3}}{n}}$$

From a social standpoint, the insurance scheme which allows for larger real projects has the higher welfare, since investors break even anyhow. Thus we can rule out the equilibrium with over-diversification if $I_{in} > I_{od}$. This reduces to:

$$\sqrt{k}\left(\sqrt{k+1} + \frac{k}{n}\left(\sqrt{k+3} - \sqrt{k+1}\right)\right) < k+1$$

From the convexity of the $\sqrt{k}$ function, we can use the following upper bound:

$$\sqrt{k+3} - \sqrt{k+1} < 2\frac{1}{\sqrt{k+1}}$$

thus it suffices to show that:

$$2\frac{\sqrt{k}}{k+1}\frac{k}{n} < \sqrt{k+1} - \sqrt{k}$$

now we can use the following lower bound:

$$\sqrt{k+1} - \sqrt{k} > 1\frac{1}{\sqrt{k+1}}$$

thus we have to show:

$$2\frac{k}{n}\sqrt{k} < \sqrt{k+1}$$

which is always true for $k \leq \frac{n}{2}$. Thus for sparse networks over-diversification is less efficient than central counterparty insurance and it is not chosen in equilibrium. □

References

Acharya, Viral and Albert Bisin, "Centralized versus Over-The-Counter Markets," *New York University, Working Paper*, 2009.

Allen, Franklin and Douglas Gale, "Optimal Financial Crises," *Journal of Finance*, 1998, *53* (4), 1245–1283.

—— **and** —— , "Financial Contagion," *Journal of Political Economy*, 2000, *108* (1), 1–33.

Babus, Ana, "Strategic Relationships in Over-the-Counter Markets," *Cambridge University, Working Paper*, 2009.

Bak, Peter, Kan Chen, José A. Scheinkman, and Michael D. Woodford, "Aggregate Fluctuations from Independent Sectoral Shocks: Self-Organized Criticality in a Model of Production and Inventory Dynamics," *NBER Working Paper, No. 4241*, December 1992.

Biais, Bruno, Thomas Mariotti, Guillaume Plantin, and Jean-Charles Rochet, "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications," *Review of Economic Studies*, 2007, *74*, 345–390.

Bolton, Patrick and David S Scharfstein, "A Theory of Predation Based on Agency Problems in Financial Contracting," *American Economic Review*, March 1990, *80* (1), 93–106.

Brunnermeier, Markus, "Deciphering the Liquidity and Credit Crunch 2007-08," *Journal of Economic Perspectives*, 2009, *23* (1), 77–100.

Caballero, Ricardo J. and Alp Simsek, "Complexity and Financial Panics," *NBER working paper, No. 14997*, 2009.

Calomiris, Charles W. and Charles M. Kahn, "The Role of Demandable Debt in Structuring Optimal Banking Arrangements," *American Economic Review*, 1991, *81* (3), 497–513.

Duffie, Darrell and Haoxiang Zhu, "Does a Central Clearing Counterparty Reduce Counterparty Risk?," *Stanford University, Working Paper*, 2009.

He, Zhiguo and Wei Xiong, "Dynamic Bank Runs," *Princeton University, Working Paper*, 2009.

Innes, R., "Limited liability and incentive contracting with ex ante action choices," *Journal of Economic Theory*, 1990, *75*, 45–67.

Kashyap, Anil K., Raghuram G. Rajan, and Jeremy C. Stein, "Rethinking Capital Regulation," *University of Chicago, Working Paper*, 2008.

Kiyotaki, Nobuhiro and John Moore, "Credit Chains," *Princeton University, Working Paper*, 1997.

Lorenzoni, Guido, "Inefficient Credit Booms," *Review of Economic Studies*, 2008, *52* (3), 809–833.

Pirrong, Craig, "The Economics of Clearing in Derivatives Markets: Netting, Asymmetric Information, and the Sharing of Default Risks Through a Central Counterparty," *University of Houston, Working Paper*, 2009.

Figures

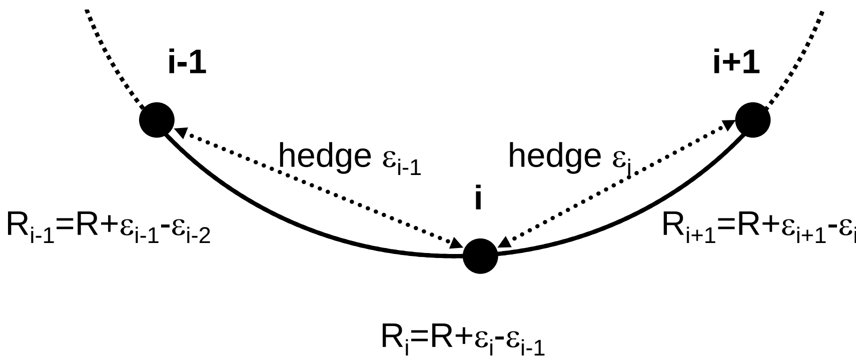

Figure 1: Market and risk structure around market i

The figure depicts the risk structure of the models in Sections 2 and 3. R_i is the return on the real project i. The ϵ shocks are offsetting between neighbors and thus can be completely hedged away.

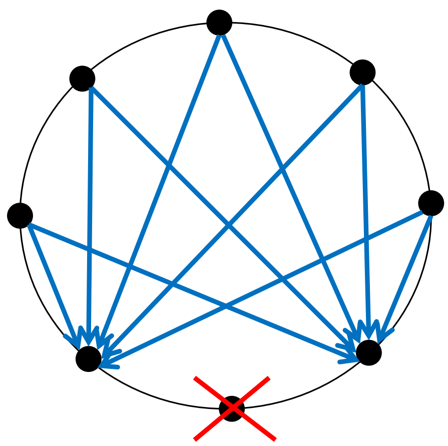

Figure 2: Socially efficient provision of equity

Capital is injected by other banks into the two banks neighboring the one that failed to ensure market stability.

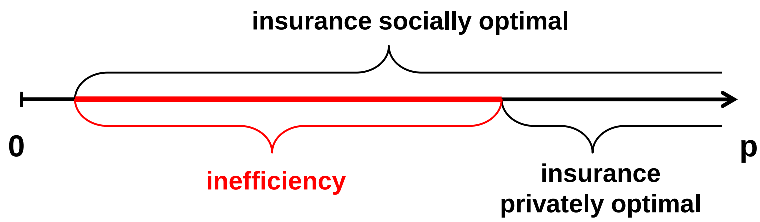

Figure 3: Efficiency and non-efficiency at different levels of p
As the probability of crisis increases, first counterparty insurance is not optimal. Above a certain threshold, it becomes socially optimal but it is not implemented in a competitive market. As the probability increases further, it even becomes privately optimal to set up a counterparty insurance scheme.

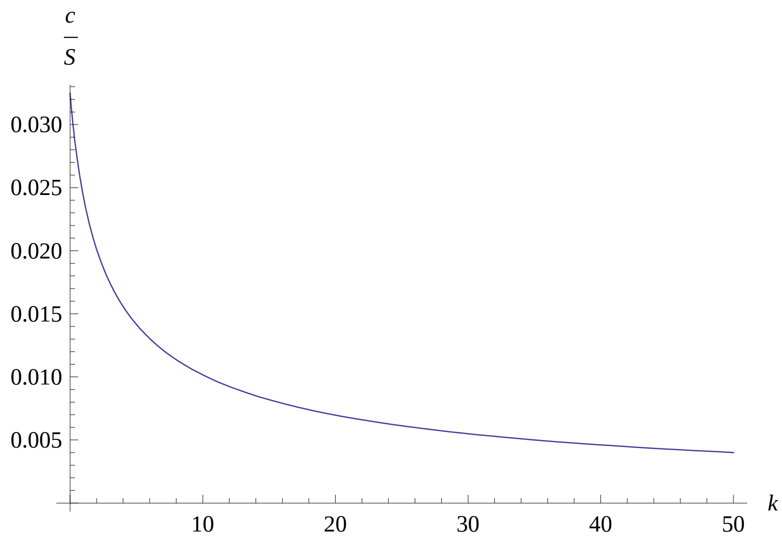

Figure 4: The density of the network and the cost of connection in general equilibrium
$\frac{c}{S}$ is the fixed cost of contracting with a counterparty as a fraction of equity, k is the number of counterparties chosen. For any k the line is the level of contracting cost for which the bank is indifferent between having k or $k+1$ counterparties. The parameters used for the graph are: $R = 1.04$, $R_f = 1$, $n = 50$, $\sigma = 0.2$, $\kappa = 2$, $p = 0.02$, $L = 0.7$.

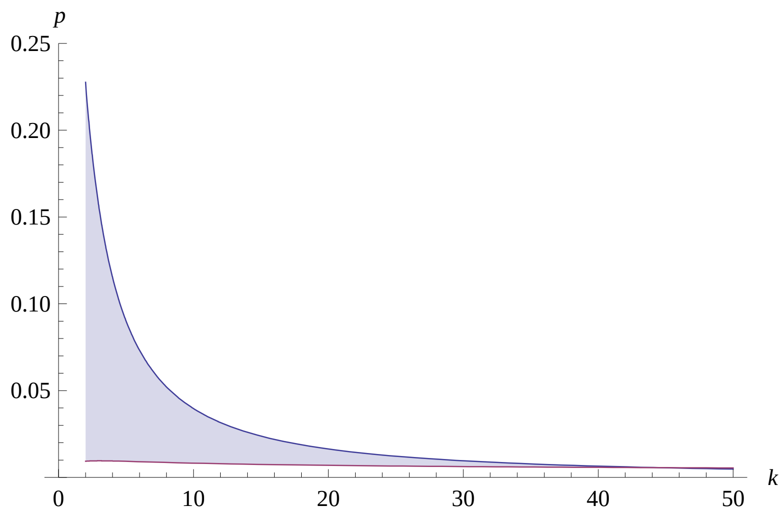

Figure 5: Inefficiency region in the general equilibrium network
p is the probability of crisis state, k is the number of counterparties. The shaded region is where counterparty insurance, even though socially optimal, cannot be sustained in equilibrium. Below the shaded region, counterparty insurance is not socially optimal; above it, it is socially optimal and can be sustained in equilibrium. The parameters used for the graph are: $R = 1.04$, $R_f = 1$, $n = 50$, $\sigma = 0.2$, $\kappa = 2$, $L = 0.7$.

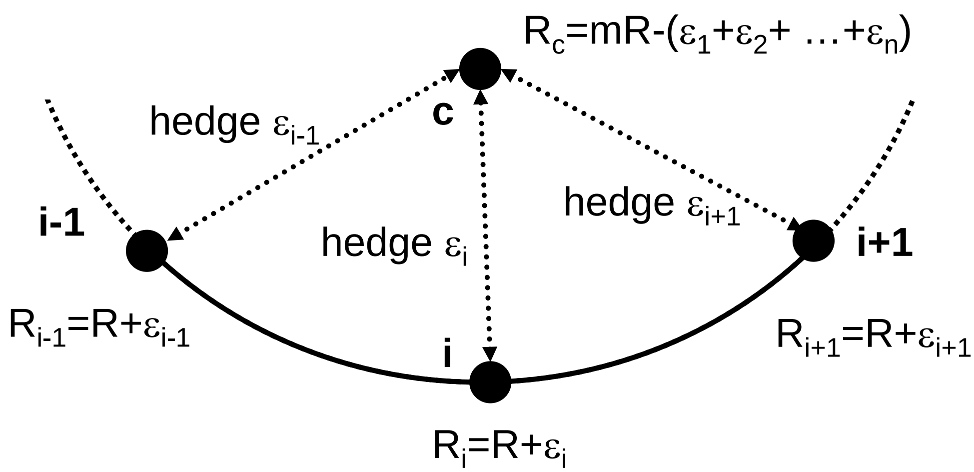

Figure 6: Market and risk structure of centralized network
R_i is the return on the real project i, R_c is the return on the real project of the central hub. The ϵ shocks are offsetting between neighbors and thus can be completely hedged away.

Tables

counterparty	**Payment channels** (billions) CDS collat.	Maiden III	total (billions)	% of total	cumulative % of total
Societe General	4.1	6.9	11.0	22.2%	
Goldman Sachs	2.5	5.6	8.1	16.4%	
Deutsche Bank	2.6	2.8	5.4	10.9%	49.5%
Merrill Lynch	1.8	3.1	4.9	9.9%	59.4%
UBS	0.8	2.5	3.3	6.7%	66.1%
Calyon	1.1	1.2	2.3	4.6%	70.7%
Barclays	0.9	0.6	1.5	3.0%	73.7%
Bank of America	0.2	0.5	0.7	1.4%	75.2%
HSBC	0.2	0.0	0.2	0.4%	75.6%
total	22.4	27.1	49.5	100.0%	

Table 1: Payments of AIG to top counterparties September 16 to December 31, 2008.
All numbers in billions of US dollars. The first column of numbers are the margin payments for credit default swaps, the second are the payments through the vehicle called Maiden III. The last column shows the cumulative payments to top counterparties.
Source: AIG disclosure, March 15, 2009. Calculated from Attachments A, B, and D.
Available online at (last accessed March 19, 2009):
http://www.aig.com/aigweb/internet/en/files/CounterpartyAttachments031809_tcm385-155645.pdf