

Entangled Financial Systems*

JOB MARKET PAPER

Adam Zawadowski
Princeton University

azawadow@princeton.edu

www.princeton.edu/~azawadow

November 16, 2009

Abstract

This paper analyzes counterparty risk in entangled financial systems. The system is “entangled” because banks hedge risks using a network of bilateral over-the-counter contracts. If banks have large exposures to a few counterparties, they do not buy insurance against a low probability counterparty default even though it is socially desirable. This is because they do not take into account that their own failure also drags down other banks — a network externality. Given that banks choose short-term financing, the failure of a single large bank prompts a systemic run. The whole system collapses even though banks have positive equity and are not directly linked through credit exposure. Mandatory counterparty insurance is welfare improving.

*I thank Markus Brunnermeier and Wei Xiong for advice, and Ana Babus, Marco Battaglini, Amil Dasgupta, Thomas Eisenbach, Harrison Hong, Paul Glasserman, Mark Flannery, Zhiguo He, Jakub Jurek, Anil Kashyap, Nobuhiro Kiyotaki, Péter Kondor, Stephen Morris, Tamás Papp, Ricardo Reis, José Scheinkman, Felipe Schwartzman, Hyun Shin, David Skeie, Ádám Szeidl, Béla Személy, Balázs Szentes, Motohiro Yogo, Tanju Yorulmazer (discussant), and conference and seminar participants at the Western Finance Association 2009 Annual Meeting (San Diego), Federal Reserve Bank of New York, Central Bank of Hungary, Princeton University macro, micro and finance workshops for useful discussions and comments on this and earlier versions of the paper. I thank the Federal Reserve Bank of New York and the Central European University for their hospitality while part of this research was undertaken. All remaining errors are mine. First draft: April 3, 2008.

1 Introduction

Modern financial institutions are entangled in a network of illiquid bilateral hedging contracts, such as over-the-counter (OTC) derivatives. The fear of these instruments affecting the whole financial system was a major argument brought up to support government intervention in the Financial Crisis of 2008.¹ In this paper, I propose a model of the core of the financial system in which institutions get endogenously entangled in OTC contracts to hedge their exposure and thus expose themselves to counterparty risk. I assess whether the failure of a single institution could lead to the collapse of the whole system. The main goal of the paper is to understand the conditions under which financial institutions have an incentive to get exposed to counterparty risk without hedging it: I show that the degree of “entanglement” in the system is crucial.

In my model, financial institutions (banks from hereon) have to satisfy risk-adjusted capital constraints. They hedge their risks using OTC contracts, which link banks together. In the event that one of the banks is hit by a large negative shock, it goes bankrupt, leaving its counterparties unhedged. Banks that have lost their hedge may not have enough capital to satisfy their capital constraint and thus could fail. If all banks are linked, the above mechanism can lead to a complete collapse of the financial system. The result that a single default can lead to a systemic crisis only holds for core banks, the financial system is resilient to the default of smaller periphery banks.

Why does the violation of the capital constraint lead to a bank’s failure? If banks are financed short-term, they need to hold enough capital to reassure investors of their incentives. In case their risk-adjusted capital is not sufficient, short-term debt is not rolled over. Thus the collapse of the system is induced by a systemic run of short-term lenders due to a crisis of confidence. If capital is costly and borrowing cheap, banks indeed 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, e.g long-term profits, as “collateral” to overcome the moral hazard problem. If a bank’s counterparty fails, it becomes riskier and its probability of default increases, thus decreasing the expected value of non-pledgable payoff. This is what makes the bank’s short-term lenders wary of its incentives, withdrawing short-term funds and leading to the liquidation of banks.

The above contagion mechanism is different from the traditional channel through direct credit exposure.² This is important since the ISDA Margin Survey 2008 reports that 65% of total OTC credit exposure is covered by collateral or margin accounts. In this paper, I show that even if there are no direct losses in assets when a counterparty defaults, contagion is possible. Even though in equilibrium there is systemic failure, it is not driven by “domino” losses in assets. The failure comes through the liability side: lenders run on banks whose hedges fail. A crucial assumption is that rehedgeing is impossible after the failure of the counterparty. This assumption captures the insight that in crisis, as in the aftermath of the Lehman bankruptcy, it is very costly or even impossible to

¹ “Wall Street’s crisis”, *The Economist*, March 22, 2008.

² See for example Allen and Gale (2000) and Kiyotaki and Moore (1997).

rehedge risks.³

The basic question is why do banks not insure against the failure of a counterparty if that could have such devastating consequences? I show that banks' individual decisions endogenously lead to a contagious network, even though this is socially suboptimal. The inefficiency comes from a network externality. Banks do not take into account that their own failure can also drag down others: their counterparties, the counterparties of their counterparties, etc. Assume for example that bank *A* is linked to bank *B*, and bank *B* is linked to bank *C*. Thus *A* is not directly linked to *C*, however, it is linked indirectly through *B*. If *B* insures against the failure of *C*, *A* can enjoy the benefits of being shielded from the failure of *C* without having to contribute and without buying insurance against the failure of *B*. Clearly, *B* does not take into account this positive externality on *A* when choosing whether to buy insurance against the failure of bank *C*. These externalities lead to a market failure: the public good of financial stability is not provided in equilibrium. This inefficient contagious equilibrium exists for relatively rare crisis events. However, if crises are very rare, it is socially optimal not to insure against them. On the other hand, if crises are very likely, then it is even individually optimal for banks to insure against counterparty default.

This paper clarifies the effects of “entanglement”: it is partially entangled financial systems, those resulting in sparse networks, that are most likely to be prone to crisis because of an inefficiently low ex ante choice of counterparty insurance. The sparse network structure of OTC contracts is crucial for the externality: it implies that every bank has exposures to only a few counterparties, so banks only care about the potential bankruptcy of their direct counterparties when making decisions. If instead the financial system was dense and each bank had the same exposure to every other bank in the system, the system would be constrained efficient. The interval of crisis probabilities, for which inefficiency prevails, shrinks as the network of hedging contracts becomes denser. The reason is that if every bank is directly linked to the failing bank, they cannot count on other banks' insurance contracts to shield them from the crisis. Since they need to have insurance of their own to survive, free riding is not possible any more. The inefficiency results regarding sparse networks extend to entangled systems where banks hedge with most or all banks in the network but have some disproportionately large exposures. Interestingly, it is exactly the largest exposures that the banks fail to insure against, since these are the most costly to insure. The limited empirical data on OTC exposures of financial institutions, such as the exposures of AIG, indicate that the financial system indeed exhibits strong counterparty exposures: see Table 1 at the end of the paper.

The externality that leads to an inefficient ex ante choice of counterparty insurance can be eliminated by regulatory intervention. One possibility is to mandate banks to buy default insurance on their counterparties from the other banks in the system. This could be implemented using a centralized insurance fund or in a decentralized way in the form of pre-funded credit default swaps (CDS). Thus this paper suggests that there are not enough CDS's in the market: a counterintuitive conclusion given that CDS contracts also link banks together leading to an entangled system. In my model,

³ “The great untangling”, *The Economist*, November 8, 2008

banks do not have enough default swaps on other banks that are their counterparties. In essence counterparty insurance makes the network dense and thus eliminates the free riding problem. It is an ex ante mechanism that stabilizes the financial system in case of the idiosyncratic failure of a bank in the core of the system, even if the regulator cannot observe the full network structure. It can be seen as a complement to the recession insurance proposal of Kashyap et al. (2008), which is designed to maintain the stability of the financial system in case of an aggregate shock affecting all banks.

An alternative policy intervention, if implementable, is to mandate the use of a central counterparty (CCP) for hedging risk. The important feature that helps overcome inefficiency in a CCP arrangement is loss mutualization: risks from the failure of a bank are spread evenly all over the system thus it makes all other banks internalize the full social cost of the collapse of another bank in the system, basically again making the network dense. Thus the reason a CCP would be welfare improving in my model is different from the reasons put forward by Duffie and Zhu (2009) and Acharya and Bisin (2009). My result that banks fail to insure against counterparty risk gives one potential explanation of why participants in the financial market have resisted the call for setting up a CCP for OTC contracts: it would basically act as an insurer, making trades more expensive.

The paper is structured as follows. Section 2 discusses how the paper relates to previous work. Section 3 shows the main inefficiency result in a baseline model with exogenous capital constraints. The full microfounded model is presented in Section 4: it endogenizes the capital constraint and highlights the role of short-term debt. Section 5 extends the analysis of the baseline model to more general network structures. In Section 6, I argue that the proposed contagion mechanism could have contributed to the Financial Crisis of 2008. Section 7 concludes and discusses avenues for further research.

2 Related literature

The seminal paper of Allen and Gale (2000) shows that while interbank deposits help banks share liquidity risk, they expose banks to asset losses if their counterparty defaults. In their setting, the probability of crisis is arbitrarily small so the crisis is socially optimal. Also, the underlying Diamond and Dybvig (1983) model does not explicitly model equity choice. This contrasts to my model where there is no direct credit exposure and I explicitly allow for anticipated crises, insurance against the default of a counterparty, and for equity choice. Allen and Gale (2000) also show that the system most prone to crisis is the one connected in a sparse network, a similar result in random networks is obtained by Gai and Kapadia (2007). However, this result is purely mechanical: if a bank has links to more banks in the interbank market, each of its interbank deposits is smaller and thus there is a smaller effect of any failure in the system. In my model, the same result is due to externalities when making ex ante choices. Babus (2009) extends the analysis to allow banks to hold mutual deposits to insure against the failure of a counterparty and finds a mixed equilibrium where some completely insure against failure; in my model, a centralized insurance market breaks down. Even

though Soramäki et al. (2007) and Bech and Atalay (2008) show that interbank lending networks are indeed sparse, empirical evidence on contagion through interbank deposits shows that it is unlikely that large fractions of the banking system could collapse due to the mechanism outlined by Allen and Gale (2000): see Furfine (2003) for evidence from the US and Upper and Worm (2004) for Germany.

Kiyotaki and Moore (1997) develop a different model where bilateral credit links result from the specificity of intermediate goods, not from hedging of liquidity risk. They point out an externality due to the chain of credit: entrepreneurs are unwilling to renegotiate loans ex post since they do not internalize the benefits accruing to others in the credit chain. However, they not allow for counterparty insurance, only for insurance against the aggregate event. Rampini and Viswanathan (2009) also show that firms with limited net worth have such a large opportunity cost of equity that they might rationally forego hedging risk, however, there is no externality in their setting.

Other papers model externalities and contagion in credit markets, but without networks: Lorenzoni (2008) shows externalities through fire-sale prices, while in Adrian and Shin (2008) and Adrian and Shin (forthcoming), the channel of contagion is due to marking-to-market accounting rules. Cifuentes et al. (2005) show that the interaction between marked-to-market pricing and networks magnifies initial shocks due to forced liquidation.

Stulz (2009) lists three positive effects of a clearinghouse in OTC markets: netting, monitoring exposures, and spreading losses. I focus on the third, while the first two effects have been modeled previously. Duffie and Zhu (2009) show a clearinghouse would help reduce costly margin accounts by netting, Brunnermeier (2009) argues that netting stops worries about the creditworthiness of a single counterparty from leading to a systemic crisis. Acharya and Bisin (2009) show that market participants build up excessive exposures to risks without monitoring total exposures. On the other hand, Pirrong (2009) argues that a clearinghouse would reduce the incentives of market participants to monitor each other.

My paper also gives an alternative explanation for the breakdown of private coinsurance arrangements between banks, such as those discussed in Calomiris (2000). Acharya et al. (2008) attribute their collapse to the market power of institutions with surplus liquidity, while my paper shows that if banks are connected in a network, coinsurance collapses due to externalities.

Research on networks in finance has uncovered a number of externalities. Rotemberg (2008) studies the amount of liquidity needed to settle interconnected financial obligations using results from graph theory. Lagunoff and Schreft (2001) show that contagion in a financial network cannot only be backward looking through asset losses, but also forward looking: agents break links to reduce their risk exposure. Leitner (2005) extends this analysis to show that links prompt luckier investors to bail-out weaker investors to prevent contagion even without any ex ante commitment, while in my paper bail-out arrangements break down because of externalities in the network. Caballero and Simsek (2009) show that uncertainty about the network structure can lead to inefficient levels of liquidity hoarding, while Zawadowski (2009b) argues that uncertainty about the availability of funding in a credit network can lead to over-hoarding of liquidity, even in the absence of any defaults in equilibrium.

3 Baseline model

I first introduce a baseline model to show the main source of inefficiency in entangled financial systems. I simply assume that there is a risk-adjusted capital constraint and banks that violate this constraint fail. Section 4 then presents a full microfounded model that rationalizes these assumptions. First, it shows how a moral hazard problem endogenously gives rise to a risk-adjusted capital constraint. Second, it proves that the capital constraint is indeed hard: banks that do not have enough capital experience a crisis of confidence and their short-term debt is not rolled over, pushing them into bankruptcy. In the generalizations of the network structure, in Section 5, I once again return to the baseline model due to tractability.

3.1 Setup of the baseline model

3.1.1 Banks

The model has three time periods: $t = 0, 1, 2$. I call these the initial, the interim and the final period. There are n markets on a circle indexed by $i = 1 \dots n$, 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 market i , bank i can use its expertise to invest a unit at $t = 0$ in a long-term real asset. This asset yields a return of $R_i = R + \epsilon_i - \epsilon_{i+1}$ in the final period, where R 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 Figure 1.

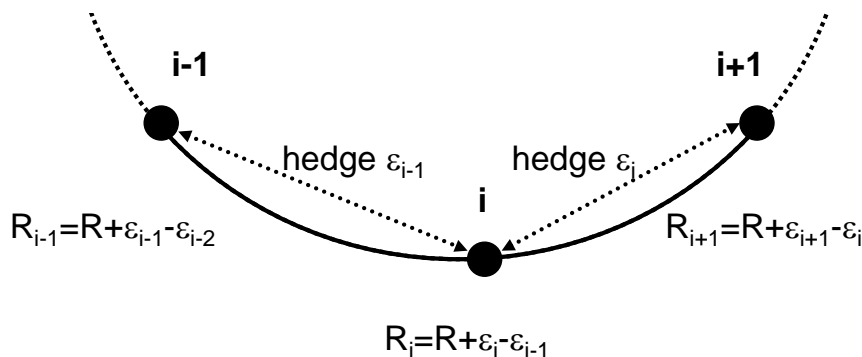


Figure 1: Market and risk structure around market i

The figure depicts the risk structure of the models in Sections 3 and 4. R_i is the return on the real project i and R is a constant. The ϵ shocks are offsetting between neighbors and thus can be completely hedged.

3.1.2 Capital requirement and payoffs

Entrepreneur i can choose to place capital K_i in the bank at $t = 0$ at an opportunity cost of c per unit. However, it is not possible to raise capital at $t = 1$, except for transferring capital from one bank to another.⁴ There is a risk-adjusted capital requirement at $t = 1$, e.g. 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 capital at $t = 1$:

$$\underline{K} = \underline{K}(\sigma_1) \tag{1}$$

where $\underline{K}(0) \geq 0$, and $\underline{K}'(x) > 0$ for $x > 0$. This capital requirement could be regulatory or endogenous and enforced by the market. If the bank does not have enough capital in $t = 1$, it goes bankrupt: I rule out partial liquidation due to the specificity of the real asset.⁵ In the full model of Section 4, I endogenize the capital constraint: investors rationally run on banks which do not have enough capital, thus the debt of these banks is not rolled over and they go bankrupt. Note that there is no explicit capital requirement as of $t = 0$ but agents rationally anticipate the capital constraint at $t = 1$ which might prompt banks to hold capital at $t = 0$. In the microfounded model the absence of an explicit capital constraint at $t = 0$ arises endogenously.

3.1.3 Uninsurable idiosyncratic shocks

There are two possible states of nature at $t = 1$. With probability p the state of nature is “crisis”: one of the banks on the circle receives an uninsurable idiosyncratic shock and goes bankrupt. The bank hit by the shock is selected with an equal probability, thus each bank fails with probability $\frac{p}{n}$. I refer to the state where there is no idiosyncratic shock, 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 a systemic crisis in general: that depends on the endogenous capital choice of the banks. In the full model of Section 4, the uninsurability of the idiosyncratic shock results from a moral hazard problem.

3.1.4 Hedging contracts and payoffs

I assume capital is costly, so banks take decisions to credibly reduce risk: they hedge their ϵ risks. I assume that only banks i and $i + 1$ can contract on ϵ_i , that is while they can sign an OTC 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 . One example for such risks are CDS contracts on complicated collateralized debt obligations. Thus the banks hold risks that can be hedged using OTC contracts, leading to a financial system entangled in

⁴A similar assumption underlies the result of Shleifer and Vishny (1997): arbitrageurs cannot raise more capital in crisis.

⁵In general it would be enough to give an upper bound on liquidation value: if liquidation is too costly, it not only reduces the balance sheet but also destroys equity.

OTC contracts. I assume that the existence of these hedging contracts is observed by all participants. Also, recall that the ϵ risks of neighboring banks are exactly offsetting: while bank i holds $+\epsilon_i$, bank $i + 1$ holds $-\epsilon_i$. This assumption is not crucial and is only introduced to simplify the analysis.

If a counterparty defaults on a hedging contract, the contract is canceled and the risk offloaded to the counterparty lands back on the bank's balance sheet. In case a bank does not have enough capital to bear this increased risk, it goes bankrupt. Even though banks cannot insure against the idiosyncratic shock that drives them into bankruptcy, they can in effect insure against the idiosyncratic shock of their counterparties by buying default insurance on them. I allow for such contracts in a centralized way: banks put up capital ex ante in an insurance fund which then pays it out to the claimants. Since banks cannot raise capital in crisis, the capital in the insurance fund has to cover all claims in all states of nature: it is pre-funded. If there is no crisis, the capital locked up in the insurance fund is returned to the banks at $t = 2$.

If a bank goes bankrupt, it is liquidated for $L < R$. Thus if a bank goes bankrupt at $t = 1$ with endogenous probability q , holds capital K , then the expected payoff of entrepreneur i is:

$$[1 - q(K_i, K_{-i}, \dots)] \cdot R + q(K_i, K_{-i}, \dots) \cdot L - 1 - c \cdot K_i \quad (2)$$

Thus the full decision problem of bank i is: choose the level of capital K_i , whether to hedge ϵ_i and ϵ_{i-1} , and what amount of default insurance to buy on its counterparties to maximize expected payoff. Note that the probability of bankruptcy depends on the capital K_{-i} and the choices of all other banks.

3.2 Systemic crisis in equilibrium

I refer to financial networks, in which the collapse of a single bank leads to the collapse of the whole system, as contagious. If in a crisis all banks fail, I call that a systemic crisis. In this subsection, I show that for some intermediate values of crisis probability p none of the banks hold enough capital or counterparty insurance to survive, even if the network is contagious.

At first let us abstract from counterparty insurance. Based on the capital requirement, the minimum amount of capital needed to survive at $t = 1$ depends on whether the bank is hedging its risks. A bank in autarchy needs to hold capital of at least $\underline{K}^a = \underline{K}(\sqrt{2}\sigma)$ since it holds two uncorrelated normal risks with standard deviation of σ . A bank with only one hedge needs $\underline{K}^u = \underline{K}(\sigma)$, while a bank with both risks hedged needs $\underline{K}^h = \underline{K}(0)$. It is clear that in normal times if banks hedged their risks with their counterparties, then they can all survive with capital $\underline{K}(0)$, since they have no risk left on their balance sheets. However, if one of the banks fails and all banks hold capital $\underline{K}(0)$, then all banks fail in crisis. To see this point, assume bank j fails exogenously in crisis. Then banks $j + 1$ and $j - 1$ both have risk σ land back on their balance sheets at $t = 1$ since they become unhedged. Given that they have only $\underline{K}(0)$ capital, they fail and now banks $j + 2$ and $j - 2$ become unhedged and thus fail, etc.

To see that holding only $\underline{K}(0)$ capital in a contagious system is a subgame perfect Nash-equilibrium,

we have to analyze the potential deviation of one bank. Note that if the system is contagious, then the only deviation that can save the bank from liquidation is to completely self-insure by holding \underline{K}^a in order to survive even if all other banks, including its counterparties, fail. In this case, the bank only fails if it is hit directly by the exogenous shock, with probability $\frac{p}{n}$. However, it has to maintain a higher level of capital. This deviation to autarchy is profitable in a contagious system if and only if the payoff in case of holding capital \underline{K}^a is higher than the payoff when holding capital of only \underline{K}^h :

$$(1-p) \cdot R + p \cdot L - 1 - \underline{K}^h \cdot c < \left(1 - \frac{p}{n}\right) \cdot R + \frac{p}{n} \cdot L - 1 - c \underline{K}^a \quad (3)$$

Thus a bank chooses to self-insure by holding substantial capital instead of hedging if and only if:

$$p > p^a = \frac{n}{n-1} \cdot c \cdot \frac{\underline{K}(\sqrt{2}\sigma) - \underline{K}(0)}{R-L} \quad (4)$$

If $p < p^a$ the contagious system is an equilibrium even if counterparty insurance is available. The reason is that in the contagious system no bank contributes to the insurance fund, thus the deviating bank has to self-insure.

Now let us assume that counterparty insurance is available and all banks participate. The insurance pool has to be large enough to increase the capital of the two counterparties of the failing bank in order to stop them from going bankrupt: $2 \cdot (\underline{K}^u - \underline{K}^h) = 2 \cdot [\underline{K}(\sigma) - \underline{K}(0)]$. Thus in essence all banks not directly affected by the crisis transfer capital to the two banks neighboring the failed bank through the insurance pool. I assume the insurance is fairly priced: given the symmetry of the setup, the private cost for every participant is to pay the costs of holding additional capital of $\frac{2}{n} \cdot [\underline{K}(\sigma) - \underline{K}(0)]$ locked up in the insurance fund.

Such an insurance scheme is sustainable if no bank chooses to opt out. Let us analyze the possible deviation of a single bank. If a bank opts out, the probability that the bank goes bankrupt and thus has to liquidate its assets increases from $\frac{p}{n}$ to $\frac{3p}{n}$, since given that it is uninsured, it also collapses if either of its counterparties collapses. The private benefit is that it does not have to contribute to the insurance fund at all. Thus a bank stays in the insurance scheme if and only if:

$$\left(1 - \frac{p}{n}\right) \cdot R + \frac{p}{n} \cdot L - 1 - \underline{K}(0) \cdot c - 2 \cdot \frac{\underline{K}(\sigma) - \underline{K}(0)}{n} \cdot c > \left(1 - \frac{3p}{n}\right) \cdot R + \frac{3p}{n} \cdot L - 1 - \underline{K}(0) \cdot c \quad (5)$$

Note that if any bank has an incentive to opt out, that means that every other bank will opt out, since they all face the same decision. Furthermore, only every other bank contributing to the scheme is not an equilibrium either: although the probability of being hit by the crisis doubles due to every other bank being uninsured, so does the price of insurance. In a similar way it follows that no mixed equilibrium is possible: see the Proof of Proposition 2 for a detailed discussion. Thus the insurance scheme is sustainable if and only if:

$$p > p_{out}^i = c \cdot \frac{\underline{K}(\sigma) - \underline{K}(0)}{R-L} \quad (6)$$

Since $p_{out}^i < p^a$, there is a range with multiple equilibria: both the systemic crisis and the stable system with insurance are equilibria. The insurance is socially optimal if the total welfare of the system with counterparty insurance is higher than that without it:

$$\begin{aligned} & (n - p) \cdot R + p \cdot L - n - n \cdot c \cdot \underline{K}(0) - c \cdot 2 [\underline{K}(\sigma) - \underline{K}(0)] \\ & > (n - n \cdot p) \cdot R + n \cdot p \cdot L - n - n \cdot c \cdot \underline{K}(0) \end{aligned} \quad (7)$$

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. Thus counterparty insurance is socially optimal if and only if:

$$p > p^s = \frac{2}{n - 1} \cdot c \cdot \frac{\underline{K}(\sigma) - \underline{K}(0)}{R - L} \quad (8)$$

Note that as $p \rightarrow 0$, the financial system completely collapses in crisis and crisis is socially optimal, as in the analysis of Allen and Gale (2000). However, in an intermediate region of crisis probabilities $p \in (p^s, p_{out}^i)$, counterparty insurance is not sustainable in equilibrium, even though it is socially optimal: I call this the inefficiency region. In Subsection 5.1, I generalize the model from two counterparties to k and show that the inefficiency region shrinks as the networks becomes more dense and eventually disappears once the network of hedging contracts is complete.

What is the missing market that makes the privately optimal choice social inefficient? In the model, I used the simplifying assumption that ϵ hedging contracts cannot be made conditional on the counterparty's decision on counterparty insurance. So called "due diligence" contracts could in theory lead to demanding that counterparties hold insurance against the default of their 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 increase their own expected payoffs. To overcome inefficiency, one needs to make sure that each contract is conditional on all counterparties in the system being insured. Thus the missing markets in my model are hedging contracts that specify that the counterparty has to hold counterparty insurance and that the counterparty has to make sure its own counterparty also holds insurance, etc. In reality it seems unreasonable that bilateral contracts requiring actions from others than the two counterparties can be written and enforced.

3.3 Possible policy responses

In case of inefficiency, mandatory counterparty insurance would be welfare improving. There are two main mechanisms to do this: using a central insurance fund like the one modeled above or a decentralized system of default swaps. The decentralized system has to insure that each bank buys the same amount of default insurance on its counterparties from all other banks. In case the sellers of the insurance have to put up the maximum amount of money they might have to pay on their insurance

contracts ex ante into a separate fund, the decentralized system is equivalent to the centralized one. The decentralized insurance scheme has the advantage of letting the market price the insurance, however, if the default swaps are not well diversified then that results in another layer of counterparty risk on top of the existing one. Having to buy default insurance on one's counterparty also discourages contracting with financial institutions with high probability of default. A potential drawback of the decentralized scheme is that if banks want to game the regulation they could easily set up a firm, like a monoline insurer, that issues very cheap default insurance for all agents to cover the needs from regulation but then fails to pay in crisis. However, this can be overcome by mandating that default insurance be pre-funded, thus capital set aside ex ante to pay in case of default events.

Another possible policy I did not discuss yet is mandating the use of a central counterparty (CCP) with loss mutualization. This would ensure that the risks after the default of a bank are spread out to all other banks, similar to that in case of mandatory counterparty insurance. However, this approach crucially assumes that the OTC contracts can be standardized and the CCP can manage them even if one of the banks who was the original counterparty goes bankrupt. The framework I use, based on special risks which can be only traded OTC, hints that such an approach might not be viable. Nevertheless, we can analyze and compare it to the results from the insurance fund. The risk in the system after the default of one bank has standard deviation of $\sqrt{2}\sigma$ and is split up between the remaining $n - 1$ banks. Thus each remaining bank needs additional capital of $\underline{K}\left(\frac{\sqrt{2}\sigma}{n-1}\right) - \underline{K}(0)$: the total amount of capital to be pledged in the so-called "guarantee fund" is $(n - 1) \cdot \left[\underline{K}\left(\frac{\sqrt{2}\sigma}{n-1}\right) - \underline{K}(0)\right]$. This can be compared with the capital of the centralized insurance fund: $2 \cdot [\underline{K}(\sigma) - \underline{K}(0)]$. For example if the capital constraint is linear $\underline{K}(x) = x$, then the amount of capital in the CCP guarantee fund is only $\sqrt{2}\sigma$, while that in the insurance fund is 2σ . Thus the CCP, if implementable, is more efficient. This is true in general if the capital requirement is convex: $\underline{K}''(x) \geq 0$. Convexity implies $\frac{n-1}{\sqrt{2}} \cdot \underline{K}\left(\frac{\sqrt{2}\sigma}{n-1}\right) < \underline{K}(\sigma)$ since $\underline{K}(0) \geq 0$ and $n > 3$ by assumption, yielding the result.

The fact that while default insurance on large banks is easily standardized, while other risks are not, leads to a potential hybrid setup. This would standardize default swaps on large banks as reference entities and make them available through the CCP, basically resulting in an insurance fund. All other OTC contracts that can be standardized can be traded through the CCP as well. For all other products that are not standardized or traded through the CCP and thus remain OTC, counterparty insurance would be mandatory. To ensure that counterparty insurance does not add another layer of entanglement to the system, they would have to be purchased from the CCP. The notional amount of insurance that would have to be bought for each OTC instrument would be determined by the amount of extra capital that would be required if the bank held it on its own balance sheet. Note that this is much more efficient than making all banks keep enough capital to survive a crisis, which in the model means forcing all banks into autarchy. The reason is that in case a bank defaults, the insurance scheme surgically injects the capital into the counterparties of the failing bank and thus stabilizes the system.

3.4 Discussion of model assumptions and robustness

In this section, I discuss some of the key assumptions of the baseline model and also argue that the main results would still hold up if some of the assumptions were relaxed. Assumptions of the microfounded model are discussed in Subsection 4.3.

A central assumption implicit in the structure of the model is that banks cannot re hedge with other banks if their counterparty fails. 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 even if counterparties could re hedge, they incurred huge losses,⁶ since they had to re hedge in a very volatile market with counterparties who were less willing to hedge away their risks. Also, many risks are specialized and thus illiquid, even more so in crisis, making them even harder to replace (Stulz 2009). The qualitative results of my analysis would hold if rehedging were possible but costly.

A theoretical question is why only two banks can contract on each ϵ risk and how these contracts are enforced. Here I propose one way to rationalize these assumptions. Assume that the realization of ϵ is observable and thus enforceable by the court. Ex ante each bank has several types of risks which are indistinguishable from the ϵ risk for any outside bank. However, not all risks have a normal distribution like the risk ϵ , they might be skewed such that if the bank can sell this risk for price zero to another bank, it makes substantial profits. Thus except for banks i and $i + 1$ which can distinguish risk ϵ_i risk from other risks, no other bank will be willing to buy risk ϵ_i , since it is afraid of “being taken for a ride”. This kind of market breakdown due to asymmetric information is akin to the lemon’s problem of Akerlof (1970).

Another important question is whether some entrepreneurs would want to sit on the sideline to buy real assets as banks fail. However, this is not the case since entrepreneurs have a large opportunity cost and crisis are relatively infrequent. Thus they would have to buy assets at huge discounts (low liquidation value) to make up for the large opportunity cost of sitting on liquid assets: liquidation in a real sense is still better.

4 Full microfounded model

This section extends the baseline model and provides a microfoundation for the contagion mechanism. It endogenizes the capital constraint and highlights the importance of short-term debt. The basic intuition is the following. When a bank becomes more risky, because of the failure of a counterparty, it is more likely to go bankrupt. This in turn means its non-pledgable payoff, that is conditional on survival, decreases and thus it cannot convince its short-term lenders that it has the proper incentives to exert effort. This leads to a crisis of confidence and short-term lenders do not roll over the bank’s debt. The section also shows that banks endogenously choose short-term financing.

A reader only interested in the insights of the paper regarding inefficiency can skip this section at

⁶ “Citadel Files Claim to Recover \$470.5 Million”, The Wall Street Journal, August 25, 2009.

first reading. The basic setup of the model is similar to that described in Section 3. Nevertheless, I still describe the full microfounded model in this subsection in order to avoid any confusion. Given the complexity of the full microfounded model, the results in this section are presented in the form of Lemmas and Propositions. All proofs are relegated to the Appendix.

4.1 Model setup

4.1.1 Participants and markets

The model has three time periods: $t = 0, 1, 2$. There are n markets on a circle indexed by $i = 1 \dots n$, where $n > 3$. There are two types of market participants: n entrepreneurs and a continuum of investors, both of them assumed to be risk-neutral. Only entrepreneurs can invest in risky real assets, investors can only invest in bank debt.

Entrepreneur i can set up a bank in market i , and no bank can operate in multiple markets. By establishing a bank, the entrepreneur becomes both manager and equityholder 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. The equity provided by the entrepreneur is the capital of the baseline model. Entrepreneurs have an outside option to invest in an asset with expected return of R_e in the long run (from $t = 0$ to 2), which they cannot liquidate 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.⁷ Thus bank i is financed by equity (stock) $K_i \geq 0$ provided by entrepreneurs and debt $D_i \geq 0$ provided by investors.

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 $t = 1$. All participants in the market have full understanding of the model of the economy, know the network structure,⁸ and act rationally.

4.1.2 Investments

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 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 like in the baseline model: see Figure 1. The new term compared to the baseline model δ_i is an idiosyncratic shock to the bank's asset value, the realization of which depends on bank i 's effort: see Subsection 4.1.3. If the investment is liquidated early at $t = 1$, it only yields $L < R$. I assume that all the liquidation value accrues to the bondholders and cannot be reinvested in the real asset.

⁷ I take it as given that they prefer simple debt securities. For a model of how debt is informationally insensitive and preferred by uninformed investors, see Gorton and Pennacchi (1990).

⁸ This assumption can be relaxed: it is enough if agents know that all banks are connected and each lender knows the counterparties of the bank it is lending to.

In addition to the real asset and bank debt, there is a third asset, an outside liquid riskless asset, I call them 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 amount of government bonds, held by the bank i is denoted by G_i , thus the accounting identity becomes $1 + G_i = K_i + D_i$.

I assume $R_e > R > R_{f,1}$.⁹ 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. However, entrepreneurs do not have any superior investment possibility in the interim period.

Projects also have a non-pledgable payoff of X : if a bank survives the last period in the model, i.e. can settle all its contractual obligations, it gets this additional payoff. It can be thought of as franchise value, 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 bank, if the bank 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).

4.1.3 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 $t = 1$. Following the model of Holmström and Tirole (1997), 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 see Subsection 3.2. If a bank goes bankrupt at $t = 1$, I assume $e_{i,1} = 1$, thus there is only private benefit 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, $\delta_i = d$, irrespective of its effort. Formally we can write:

$$\delta_i = \max [d(1 - e_{i,0}), d(1 - e_{i,1}), \mu_i] \tag{9}$$

⁹ The required return on equity is taken as exogenous here but could be endogenized using scalable investments and limited entrepreneurial capital, see Zawadowski (2009a).

for all $i = 1 \dots n$. δ_i is stochastic since it depends on a stochastic variable μ_i , which I call adverse shock:

$$\mu \sim \begin{cases} \text{with probability} & 1 - p : \quad \forall i : \mu_i = 0 \\ \text{with probability} & p : \quad \forall i \neq j : \mu_j = 0 \\ & \text{for } i = j : \mu_i = d \end{cases} \quad (10)$$

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

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, $\delta_i = 0$ if bank i exerts full effort, and $\delta_i = d$ if it shirks in either period. However, in times of crisis, 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 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:

$$s_i = \max[d(1 - e_{i,0}), \mu_i] \quad (11)$$

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 μ . Since the effort choices are only observable by bank i , outsiders cannot tell apart bad luck from shirking. In equilibrium, bad realizations of s_i are due to bad luck but any insurance leads to a moral hazard problem, making these idiosyncratic risks non-insurable: see Lemma 7. 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.

4.1.4 Contracts

Banks can choose between long-term and short-term debt financing. Short-term debt has to be rolled over at $t = 1$, thus debtholders have an option to withdraw funding and force the bank to liquidate the 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 in the initial period. 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 $t = 0$. 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. Given that large financial institutions usually issue bonds and commercial papers, renegotiation does not seem to be a viable option in practice.

As in the baseline model, banks can write ϵ hedging contracts. To establish an ϵ contract between

two banks, both banks have to agree in the initial period to enter into the contract. I rule out the Nash equilibrium of no hedging contracts if there is also an equilibrium where the contract is established: in essence I use the pairwise stability concept of Jackson and Wolinsky (1996). Why do banks choose to hedge at all, since debt contracts induce risk-shifting? The reason is that in equilibrium the banks choose short-term debt, and the riskier the bank, the more equity it has to keep in order to roll over its debt.

I also allow for buying counterparty insurance from a centralized insurance fund. Like in the baseline model, banks have to put up some of their equity ex ante since the fund has to be able to pay out all claimants in crisis. For simplicity, I assume that the price of ϵ hedging contracts cannot be made conditional on the counterparty having default insurance on its own counterparties. In practice it is indeed hard to verify counterparty insurance, since it can easily be undone by offsetting contracts. I discuss in Subsection 4.3 the importance of this assumption. 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.¹⁰ Given that the expected value of every ϵ is zero at $t = 1$, this is their fair value. However, if banks survive until the final period once ϵ risks are realized, then they first have to settle their ϵ hedging contracts before paying back debt. These assumptions model the main features of U.S. bankruptcy laws regarding derivatives (Bliss and Kaufman 2006).

4.1.5 Timeline of events and information sets

The timeline of events is the following. At $t = 0$, the initial period, bank i , using its own resources, decides on whether to invest one unit in real asset i and how much government bonds G_i to hold. Then it can decide whether to leverage up using either long-term or short-term debt: one way to do this is through share repurchases. It then chooses effort level $e_{i,0} \in \{0, 1\}$ and whether to sign OTC 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 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 because its short-term financing is not being rolled over, 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,

¹⁰Counterparties can close-out contracts with a failing bank and demand the replacement value. However, if the replacement value is not covered by collateral the counterparty becomes a general creditor and is subject to the automatic stay provision in bankruptcy. Thus unless the liquidation value is high, counterparties suffer substantial losses or have to take back the risk on their balance sheets.

get an extra non-pledgable payoff of X . Finally, all participants consume their payoffs.

4.1.6 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. Given that in Subsection 4.2.4 I show that in equilibrium long-term debt financing is not offered by investors due to the commitment problems in the model, here I only present the choice in case of short-term debt.

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

$$(1 - \pi_1) \cdot E_{t=0} [\max(0, R_i + R_{f,1}G_i - R_{l,1}R_{l,0}D_i + C_i)|A] + (1 - \pi_1) \cdot (1 - \pi_{2|1}) \cdot X - R_e \cdot K_i \quad (12)$$

where the event A is survival at $t = 1$. The maximization is subject to the following constraints:

$$(1 - \pi_1) \cdot E_{t=0} [\max(0, R_i + R_{f,1}G_i - R_{l,1}R_{l,0}D_i + C_i)|A] + (1 - \pi_1) \cdot (1 - \pi_{2|1}) \cdot X - R_e \cdot K_i \geq 0 \quad (13)$$

$$1 + G_i = K_i + D_i \quad (14)$$

$$E_{t=1} [\max(0, R_i + R_{f,1}G_i - R_{l,1}R_{l,0}D_i + C_i)] + (1 - \pi_{2|1}) \cdot X \geq B_1 \quad (15)$$

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 $t = 1$ conditional on surviving. 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 $t = 1$. $\pi_{2|1}$ is the probability that the bank fails at $t = 2$, conditional on surviving at $t = 1$. $R_{l,0}$ is the endogenously determined interest rate paid by the entrepreneur on its short-term debt from $t = 0$ to 1, while $R_{l,1}$ is that from $t = 1$ to 2. 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, this occurs with probability $(1 - \pi_1) \cdot (1 - \pi_{2|1})$. 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 incentive compatibility constraint at $t = 0$ is not spelled out since it is satisfied in equilibrium, otherwise investors would not lend in the first place.

4.1.7 Debtholder's choice

Interest rates are determined endogenously such that investors break even. Investors anticipate the banks' equity and investment choices in equilibrium. 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 share equally. This 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. In general, even if the model has a good equilibrium there is a bank-run equilibrium. Following Allen and Gale (2000), I rule out the bank-run equilibrium in case there is a non bank-run equilibrium as well.

I assume that debt does not carry any covenants, thus the amount of debt and the interest rate on it cannot be contingent on any given level of equity, thus the bank takes interest rates as given. However, ex post the decision whether to roll over debt can depend on the amount of equity held by the bank. This is an important simplifying assumption, since equityholders can take interest rates as given when making decisions. Also, this assumption 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. This assumption is by no means central to the results: see the discussion in Subsection 4.3.

4.1.8 Parameter restrictions and assumptions

A following parameter restrictions are needed in order to make the problem relevant to modeling counterparty risk in a contagious network:

$$L > R - d > 0 \tag{16}$$

$$d > B_0 + B_1 \tag{17}$$

$$B_1 \geq (R - R_{f,1}) + X \tag{18}$$

$$R - B_1 > R_{f,1}L \tag{19}$$

The restriction of Equation 16 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 between investors. Restriction 17 means that it is socially optimal to exert effort. Equation 18 ensures that banks have to keep at least some equity to overcome moral hazard and banks cannot operate with negative equity. Restriction 19 implies that a bank does bankrupt if its debt is not rolled over in the interim period.

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 make sure 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 1 does both. Note that the coefficient of R_e is positive (by restriction 18), making Assumption 1 an upper bound on R_e .

Assumption 1.

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

The following technical Assumption 2 is made to facilitate the solution of the model. It is needed because in the model hedged banks hold no risk, thus there is no risk-return tradeoff. In a more general model where banks can choose how much risk to keep on their balance sheets and there is a risk-return tradeoff, this assumption could be relaxed. However such a model would be much less tractable without providing much additional insight.

Assumption 2.

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

4.2 Systemic run and crisis in equilibrium

4.2.1 Equilibrium with contagion

First, let us abstract from counterparty insurance, it is added to the analysis in Subsection 4.2.6. Given the parameter restriction of Subsections 4.1.8, I prove the existence of the following subgame perfect Nash equilibrium for an intermediate range of crisis probabilities. 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 hold equity that is just enough to make the incentive constraint to exert effort in normal times satisfied at $t = 1$. This means there is no slack equity to use in crisis. When one of the banks defaults, all banks default. The defaults are driven by all short-term debtholders denying to roll over debt. Debtholders run since they know that if the neighboring bank collapses and they do not run then the bank they lent to becomes risky because it lost one of its ϵ hedges. This increases the probability that it ends up bankrupt at $t = 2$ and in turn reduces the expected value of the non-pledgable payoff as of $t = 1$. 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 the debtholders to withdraw their funding and liquidate the project instead of letting it mature while the bank shirks.

In the following I show, step by step, that the proposed equilibrium exists: this yields the micro-foundation for the baseline model of Section 3. 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 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.

4.2.2 Minimal equity needed for rolling over debt

Lemma 1. *Banks do not hold government bonds: $G_i = 0$ for all i .*

The basic insight is that holding government bonds does not alleviate the incentive constraint since it does not mean having more equity. Since government bonds have a lower return than that at which banks can borrow, they have no incentives to hold any.

In the following I set $G = 0$, since by Lemma 1 holding government bonds is never optimal for a bank. I conjecture that the banks use short-term debt and that they hedge their ϵ risks. I verify this in Lemmas 5 and 6. It is also assumed that in case of bankruptcy, shareholders are wiped out completely, a conjecture that is proved in Lemma 3. For notational simplicity I drop the subscript i since entrepreneurs are ex ante symmetric.

Now we calculate the minimal amount of equity $\underline{K}(R_{l,0}, \sigma_1)$, 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 $t = 1$. Banks need to hold equity since creditors are worried that banks could misbehave in the interim period and collect private benefits B_1 .

Lemma 2. *If a bank has a final payoff with expected value R and normally distributed risk σ_1 as of $t = 1$ and is financed by short-term debt, then 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:*

$$\underline{K}(R_{l,0}, \sigma_1) = \frac{B_1 - (R - R_{f,1}R_{l,0}) - [1 - \pi_{2|1}(R_{l,0}, \sigma_1)] \cdot X}{R_{f,1}R_{l,0}} \quad (20)$$

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

$$\pi_{2|1}(R_{l,0}, \sigma_1) = \Phi \left(\frac{R_{l,1}R_{l,0}[1 - \underline{K}(R_{l,0}, \sigma_1)] - R}{\sigma_1} \right) \quad (21)$$

and Φ denotes the cumulative density function of the standard normal distribution. The short-term rate $R_{l,1}$ charged to the bank at $t = 1$ is determined endogenously. Also, $\underline{K}(R_{l,0}, \sigma_1) \geq 0$ and $\frac{\partial \underline{K}}{\partial \sigma_1} > 0$.

This lemma provides the microfoundation for the capital constraint in the baseline model, since the function $\underline{K}(x) = \underline{K}(R_{l,0}, x)$ fulfills the two assumption made there: $\underline{K}(0) \geq 0$ and $\underline{K}'(x) > 0$. Also note that the minimal amount of equity \underline{K} is decreasing in the bank's expected non-pledgable payoff $(1 - \pi_{2|1}) \cdot X$. This is important since it highlights that non-pledgable payoff is crucial: it is used as "collateral" to overcome the moral hazard problem. Thus if the expected value of non-pledgable payoff decreases because of an increase in the probability of bankruptcy $\pi_{2|1}$, 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.

4.2.3 Minimal equity is not enough in crisis

To show that all banks collapse in crisis if they all hold equity of only $\underline{K}^h(R_{l,0}) = \underline{K}(R_{l,0}, 0)$, as if they were riskless, two steps are needed. First, in Lemma 3, I show that the bank, the market of which is hit in crisis by the adverse μ shock, fails. As a next step, in Lemma 4, I show that if bank i collapses and leaves its counterparty, bank $i + 1$, unhedged, that counterparty will also suffer a run of creditors, i.e. its debt will not be rolled over and thus it too will fail. Since all banks are linked, lenders of all banks will rationally run.

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

The interpretation is straightforward. Consider the bank, the long-term project of which is guaranteed to have a bad payoff, either because of the shock μ 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 rationale 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: see Subsection 4.3 for a discussion.

The next step, formalized in Lemma 4, is at the heart of the contagion mechanism: if all banks hold only an equity of $\underline{K}^h(R_{l,0})$, the failure of a counterparty, 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 by a run.

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

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 counterparty 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.

4.2.4 Choice of debt maturity and hedging

Now I show in Lemmas 5 and 6 that ϵ -hedging and short-term borrowing is indeed 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.

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. Furthermore, 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. Thus in this model, the short-term debt is an optimal outcome like in Calomiris and Kahn (1991) and Diamond and Rajan (2001). In practice it might be the case that using short-term debt is an inefficient outcome like in Brunnermeier and Oehmke (2009), however, since I focus on the inefficiencies due to the network of hedging contracts, I choose to model short-term debt as an efficient choice.

Lemma 6. *For any $i \in \{1 \dots n\}$, 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.*

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 OTC 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 also makes clear that banks do not simply engage in hedging to lower the capital requirement set by the regulator: there is an underlying moral hazard issue that is the "raison d'être" of these contracts.

4.2.5 Contagious system in equilibrium

Similarly as in the baseline model, a potential deviation from the proposed contagious Nash-equilibrium is that a bank chooses to hold a higher amount of equity, to make sure it survives the crisis. An autarchic bank holds total risk of $\sqrt{2}\sigma$, so the amount of equity it needs is: $\underline{K}^a(R_{l,0}) = \underline{K}(R_{l,0}, \sqrt{2}\sigma)$. 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, no bank chooses to self-insure. Proposition 1 formalizes this intuition and gives a cut-off value in the probability of crisis p .

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

$$\hat{p}^a = \frac{R_e[\underline{K}^a(R_{l,0}^h) - \underline{K}^h(R_{l,0}^h)] - \tilde{\pi}_{2|1}X}{\frac{n-1}{n}B_1 - \tilde{\pi}_{2|1}X} \quad (22)$$

The equation is implicit for the cutoff value of \hat{p}^a since both $\underline{K}^a(R_{l,0}^h)$ and $\underline{K}^a(R_{l,0}^h)$ depend on \hat{p}^a through $R_{l,0}^h$. $\hat{\pi}_{2|1}$ 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:

$$R_{l,0}^h = \frac{1}{1 - p \left(1 - \frac{R_{f,1}L}{R+X-B_1} \right)} \quad (23)$$

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

4.2.6 Suboptimal counterparty insurance

Since the adverse idiosyncratic shock μ 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. This also means that any policy which saves failed banks to prevent systemic crisis creates a moral hazard problem, thus relying on ex-post bailouts is not a reasonable policy in this model.

Lemma 7. *It is not feasible to have an insurance scheme that aims at saving the bank hit directly by the adverse idiosyncratic shock in crisis.*

However, insuring against the default of another bank is feasible, since it does not entail the same moral hazard problem. Under such a scheme, the banks neighboring the one that failed get a state contingent transfer to raise their equity and allow them to roll over debt. Since the counterparty insurance fund in place has to be fully funded ex ante, it has to be of the size of $2 \cdot [\underline{K}^u(R_{l,0}) - \underline{K}^h(R_{l,0})]$, similarly to that in the baseline model. Proposition 2 contains the main results of the paper: in an intermediate range of crisis probabilities p , counterparty insurance cannot exist in equilibrium even though it is socially optimal. Thus the inefficiency result of the simple baseline model of Section 3 holds up in the full microfounded model.¹¹

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

- (i) *if $p < \hat{p}^s$, it is not socially optimal to insure against counterparty risk*
- (ii) *if $\hat{p}^s \leq p < \hat{p}^i$, counterparty insurance is socially optimal but it cannot be sustained in equilibrium*
- (iii) *if $p \geq \hat{p}^i$, the stable equilibrium with voluntary counterparty insurance is sustainable*

The implicit equations for \hat{p}^s and \hat{p}^i are:

$$\hat{p}^s : \quad p = \frac{2}{n-1} \cdot \frac{R_e - R_{f,1}}{R_{f,1} \hat{R}_{l,0}} \cdot \frac{\hat{\pi}_{2|1} X}{R + X - R_{f,1} L - \frac{2}{n-1} \hat{\pi}_{2|1} X} \quad (24)$$

$$\hat{p}^i : \quad p = \frac{R_e - R_{f,1}}{R_{f,1} \hat{R}_{l,0}} \cdot \frac{\hat{\pi}_{2|1} X}{B_1 - \hat{\pi}_{2|1} X} \quad (25)$$

¹¹ Proving $\hat{p}^i < \hat{p}^a$ is tedious and does not provide any additional insights so it is not included in this theorem.

the equations are implicit since $\hat{R}_{t,0}$, the endogenous interest rate at $t = 0$, depends on p . $\hat{\pi}_{2|1}$ is the endogenous probability that a bank goes bankrupt at $t = 2$ if it is the counterparty of a failed bank and got ample capital infusion through the insurance scheme to roll over its debt.

4.2.7 A numerical example

For illustrational purposes I give a numerical example for which all the parameter restrictions and assumptions are satisfied and the unique equilibrium is the one with systemic crisis. I use $R = 1.01$, $R_{f,1} = 1$ to model that the assets of banks are not much more profitable than their liabilities. I set $n = 15$ to model the core of the global financial system. 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 $L = 0.7$, 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 $\sigma = 0.1$. 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 and the average annual return on common equity for investment banks was 16% before the crisis.¹² The probability of crisis in the baseline model is chosen to be $p = 0.05$, corresponding to crisis few times a century. The minimal equity share of a bank in the contagious equilibrium is $\underline{K}^h = 0.053$, which means the banks in the model work with about 20 leverage, which was not unusual for banks just before the crisis of 2008.

4.3 Discussion of model assumptions and robustness

In this section, I discuss some of the key assumptions of the microfounded model and also argue that the main results would still hold if some of the assumptions were relaxed.

The choice of using a private benefit framework instead of risk-shifting is due to two reasons. 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). Second, if one looks at recent years, there is evidence for banks engaging in activities that can be modeled by private benefits: such as accounting profits being pumped out of a company in the form of dividends, equity repurchases and high bonuses before the long-term projects mature. For example, Acharya et al. (2009) show that some banks kept on paying out substantial dividends during the Financial Crisis of 2007-2008, irrespective of large losses and dwindling capital.

In the baseline model, debt does not have any covenants, thus debtholders cannot ex ante force the bank to hold a given level of equity or to sign a given contract. This assumption is used for two reasons. First, bondholders indeed have limited influence on the financing and investment choices of a large financial institution: such complete contracts are very hard to write. Second, to maintain the simplicity of the analysis: one can then ignore the effect of the firm's decisions on the interest rate it is charged by debtholders in the initial period. Even in the presence of covenants on equity choice,

¹² "Capital spenders" in "A special report on international banking", The Economist, May 19th, 2007.

if the temptation to misbehave in the initial period B_0 is high enough, banks still choose short-term financing and the system is contagious.

The model assumes a rational run of lenders in order to emphasize that the results do not hinge on inefficient runs. Clearly the results would still hold if we allowed for inefficient runs. One possibility would be to use the dynamic model of He and Xiong (2009) where investors holding the shortest term debt run if the fundamental of a bank deteriorates or it becomes too risky, even if the bank itself is solvent.

An important assumption is that in crisis short-term bondholders get the value of the liquidated real project and are not liable towards the counterparties of the failed bank. A potential alternative setup would be one in which bondholders of the failed institution are forced to pay their counterparties some compensation for the broken contract 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. However, such a setup is not realistic for several reasons. First, it is not possible under current U.S. bankruptcy laws (Bliss and Kaufman 2006). Second, problems with implementing such a debt contract are similar to that in the case of renegotiation: the main benefit of a debt contract is its simplicity. Third, in a more general setting, such debt contracts give incentives to banks to contract with weak counterparties and leads to reduced market monitoring (Bergman et al. 2003).

5 Extensions to more general network structures

5.1 The density of networks

The baseline framework can be easily extended to incorporate more counterparties than two. Here I assume all parts of the model in Section 3 are unchanged, except for the risk structure. Each bank has exactly $k \in \{0, 1, 2, \dots, n - 1\}$ counterparties if the return on the real return of e.g. bank 1 is the following:

$$R_1 = R + \sum_{l=2}^{k+1} \epsilon_{1,l} \quad (26)$$

and bank l holds the exactly offsetting risk of $-\epsilon_{1,l}$. As in the case of the circle, I assume that only banks 1 and l can contract on risk $\epsilon_{1,l}$. In the same way, all banks have exactly k natural counterparties. The offsetting risks are realized at $t = 2$ and are i.i.d. $N(0, \sqrt{\frac{2}{k}}\sigma)$. Thus an unhedged bank holds risk with standard deviation of $\sqrt{2}\sigma$: the amount of risk of each bank does not change as the number of counterparties changes. The resulting hedging networks with equal degree (number of links) at each node can be represented by a so called regular graph. Two additional assumptions are needed. First, n is even: this ensures the existence of regular graphs with any number of counterparties k . Second,

I assume the network is fully connected if $k \geq 2$, thus all banks are connected with all other banks either directly or indirectly. See Figure 2 for an example. Note that in the special case of $k = 2$ we arrive at the circle model of Section 3.

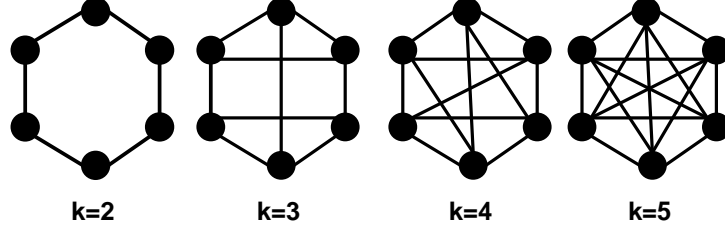


Figure 2: Varying density networks

A graph of the core of the financial system analyzed in Subsection 5.1 for $n = 6$ and different number of counterparties k .

The solution closely follows that in case of a circle, I only highlight the differences. In case a bank chooses autarchy, it has to hold exactly the same amount of capital as in case of the circle, thus the contagious equilibrium exists if Equation 4 holds, irrespective of the number of counterparties. A bank decides to opt out from counterparty insurance if and only if the expected payoff without insurance is higher than that with insurance. The size of the insurance fund is $k \cdot \left[\underline{K} \left(\sqrt{\frac{2}{k}} \sigma \right) - \underline{K}(0) \right]$ since any failing bank has k counterparties. If a bank opts out while the other banks stay insured, it goes bankrupt if it or one of its counterparties is directly hit by the crisis, i.e. with probability $\frac{(k+1)}{n} \cdot p$:

$$\begin{aligned} & \left(1 - \frac{p}{n}\right) \cdot R + \frac{p}{n} \cdot L - 1 - \underline{K}(0) \cdot c - k \cdot \frac{\underline{K} \left(\sqrt{\frac{2}{k}} \sigma \right) - \underline{K}(0)}{n} \cdot c \\ & > \left(1 - \frac{(k+1) \cdot p}{n}\right) \cdot R + \frac{(k+1) \cdot p}{n} \cdot L - 1 - \underline{K}(0) \cdot c \end{aligned} \quad (27)$$

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. Thus the insurance equilibrium is sustainable in equilibrium if and only if:

$$p < p_{out}^i(k) = c \cdot \frac{\underline{K} \left(\sqrt{\frac{2}{k}} \sigma \right) - \underline{K}(0)}{R - L} \quad (28)$$

On the other hand, the social planner will choose insurance if and only if:

$$\begin{aligned} & (n - p) \cdot R + p \cdot L - n - n \cdot c \cdot \underline{K}(0) - c \cdot k \cdot \frac{\underline{K} \left(\sqrt{\frac{2}{k}} \sigma \right) - \underline{K}(0)}{n} \\ & > (n - n \cdot p) \cdot R + n \cdot p \cdot L - n - n \cdot c \cdot \underline{K}(0) \end{aligned} \quad (29)$$

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. Thus for $k \geq 2$, counterparty insurance is socially optimal if and only if:

$$p > p^s(k) = \frac{k}{n-1} \cdot c \cdot \frac{\underline{K}\left(\sqrt{\frac{2}{k}}\sigma\right) - \underline{K}(0)}{R-L} \quad (30)$$

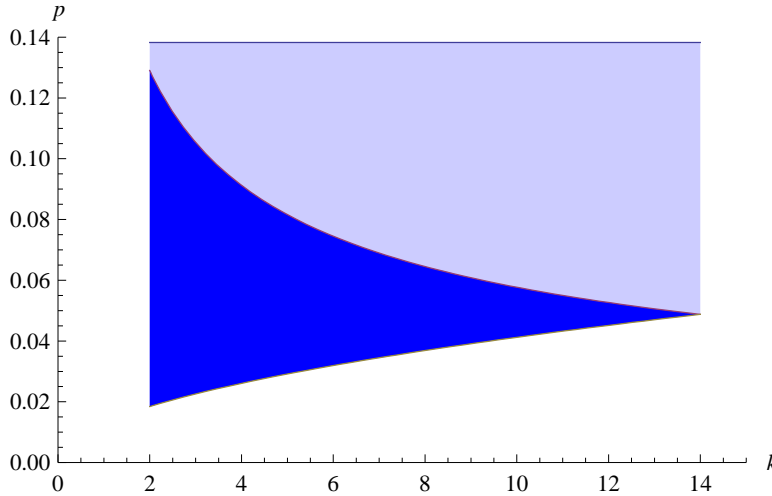


Figure 3: Inefficiency region in networks of different densities

The vertical axis is the probability of crisis state p , the horizontal axis is the number of counterparties k . The dark shaded region is where counterparty insurance, even though socially optimal, cannot be sustained in equilibrium. In the light shaded region, both the contagious system and the stable one with counterparty insurance are equilibria. Below the shaded regions, counterparty insurance is not socially optimal. Above the shaded region, counterparty insurance is sustainable in equilibrium, while the contagious equilibrium does not exist. The figure is based on the analysis of Subsection 5.1, the capital requirement is $\underline{K}(x) = x$, and the parameters used are: $c = 0.2$, $\sigma = 0.2$, $R = 1.01$, $L = 0.7$, $n = 15$.

Figure 3 shows the regions of different equilibria for the capital constraint of $\underline{K}(x) = x$. The parameters are calibrated similarly to Subsection 4.2.7. The interval of crisis probabilities for which inefficiency prevails is the largest for sparse networks of OTC hedges. The figure shows that the network can even be considered sparse, and thus likely to be inefficient, when each bank connects to a quarter or half of all banks in the system. On the other hand, for completely dense networks with $k = n - 1$, where all banks are connected to all other with the same size of hedging contracts (represented by a complete graph): $p^s(n - 1) = p_{out}^i(n - 1)$ and there is no inefficiency. Note that multiple equilibria are still present even if all banks are linked to all others, since $p^s(n - 1) < p^a(n - 1)$.

Why does the inefficiency region shrink as the network becomes denser? The intuition is that in sparse networks banks can free-ride on the counterparty insurance of others, while in dense networks they cannot. Assume for example that bank A is linked to bank B , and bank B is linked to bank

C , however, A is not directly linked to C . If B insures against the failure of C , A can enjoy the benefits from being shielded from the failure of C without having to contribute and without buying insurance against the failure of B . However, if bank A is also directly linked to bank C , then it has to buy counterparty insurance itself on bank C , it cannot simply free-ride on the insurance of bank B . Note that in a completely dense network where all banks are connected to all other banks, every bank has to buy counterparty insurance on all the other banks. All banks face the same decision and in equilibrium no bank can free-ride on another bank's counterparty insurance: the externality disappears. However, in Subsection 5.2 I show that this argument does not go through for completely dense networks where some of the links are stronger than others, even though all banks are connected with all others.

5.2 Heterogenous counterparty exposures

While the assumption that banks have large OTC contracts with some banks while no exposure to others captures some aspect of the financial system, it is very simplistic. See for the example the exposures of AIG in Table 1 or the network of interbank lending in Soramäki et al. (2007). In general, banks seem to have some disproportionately large exposures within the core of the financial system, while they have smaller but not zero exposure to others. In this section, I show that the main results of the paper carry through to such a case. Simply that banks in the core hedge with all other banks does not mean that the network is dense and thus efficient in the sense of Subsection 5.1.

Instead of a circle of exposures, I assume each bank has two strong links (large exposures) to two counterparties, such as in the baseline model of Section 3. The strong links are arranged in a circle as before. Banks are also exposed to all other banks through weak links (small exposures): see Figure 4. All banks are the same size, thus they are ex ante symmetric. Strong links are based on normally distributed offsetting risks of standard deviation σ_s , while weak links on offsetting risks of standard deviation $\sigma_w \leq \sigma_s$. The overall risk held by each bank in autarchy is assumed to be the same as before, thus $\sqrt{2 \cdot \sigma_s^2 + (n-3) \cdot \sigma_w^2} = \sqrt{2}\sigma$. All risks are uncorrelated and the rest of the setup is the same as in the baseline model. I assume that banks can separately decide whether or not to insure against each of their counterparties.

The crisis probability at which at least one bank deviates to autarchy is the same as in Equation 4. In a full insurance equilibrium, the insurance fund has to hold enough capital to stabilize the two strongly linked counterparties and the $n-3$ weakly linked counterparties of the failed bank, thus altogether: $2 \cdot [\underline{K}(\sigma_s) - \underline{K}(0)] + (n-3) \cdot [\underline{K}(\sigma_w) - \underline{K}(0)]$. Thus the insurance fund consists of two distinct parts: insurance for strong links and insurance for weak ones. Banks now have several possible ways of opting out of the insurance: they can stop buying insurance on strong links, weak links, or both. The main insight is that, in case of full insurance, opting out from insurance on any counterparty exposes the bank to the same probability of default by assumption. This is because each counterparty, irrespective of the strength of the link, has the same probability $\frac{p}{n}$ of exogenous default. Thus a deviating bank will first choose to quit insurance on its strong links, since these are

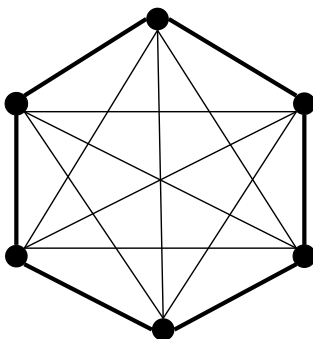


Figure 4: Heterogenous links within the core

A stylized graph of the core of the financial system analyzed in Subsection 5.2 with $n = 6$. Thick lines represent strong links (large counterparty exposure), while thin lines weak links (small counterparty exposure).

more costly than insuring weak links. The condition for deviation is the same as in case of the circle. Simply replacing σ with σ_s in Equation 5, the insurance scheme is sustainable if and only if:

$$p > \tilde{p}_{out}^i = c \cdot \frac{\underline{K}(\sigma_s) - \underline{K}(0)}{R - L} = c \cdot \frac{\underline{K}\left(\sigma \cdot \sqrt{1 - \frac{n-3}{2} \cdot \frac{\sigma_w^2}{\sigma^2}}\right) - \underline{K}(0)}{R - L} \quad (31)$$

In the special case where $\sigma_w = 0$ we arrive at Equation 6 which is the expression for the baseline model with the circle. On the other hand if $\sigma_w = \sqrt{\frac{2}{n-1}}\sigma$, thus $\sigma_w = \sigma_s$, we arrive at Equation 28 describing the completely dense network of $k = n - 1$. Note that once banks give up their insurance on strong links, the system is contagious and all banks fail in crisis. This implies that insuring weak links is useless, thus once the insurance on strong ones is dropped, so are the ones on weak links.

The crisis probability at which counterparty insurance becomes socially optimal is calculated exactly as in Equation 7 by replacing the size of the insurance pool, which in case of a circle was $2 \cdot [\underline{K}(\sigma) - \underline{K}(0)]$. Thus insurance is socially optimal if and only if:

$$p > \tilde{p}^s = c \cdot \frac{\frac{2}{n-1} \cdot [\underline{K}(\sigma_s) - \underline{K}(0)] + \frac{n-3}{n-1} \cdot [\underline{K}(\sigma_w) - \underline{K}(0)]}{R - L} \quad (32)$$

Given that $\underline{K}'(x) > 0$, in the intermediate range of $\sigma_w \in (0, \sqrt{\frac{2}{n-1}}\sigma)$: $\tilde{p}^s < \tilde{p}_{out}^i$, thus there is still an interval of crisis probabilities with inefficiency which shrinks as the weak links become closer in size to the strong links. The intuition is simply that if the weak links are small compared to the large links, the system in essence still looks like a circle and the externalities lead to a breakdown of the insurance scheme. However, as the weak links approach the size of strong links, the network in essence becomes one like the completely dense network which is efficient.

5.3 Beyond the core of the financial system

All networks considered up to now showed that the failure of a single counterparty is enough to drive the whole system into a systemic default. In a more general setting with heterogenous banks some large banks in the core of the system might hold enough capital to survive the default of small counterparties. A simple case is analyzed in this section and I show that while banks in the core of the financial system do prepare for the collapse of small periphery banks they are connected to, they do not prepare for the collapse of a large bank in the core.

I consider a simple extension of the baseline model of Section 3. There are two types of banks. Core banks have investment size of 1 and are linked in a circle, as in the baseline, with offsetting risks of standard deviation σ_c . There are $n \cdot k_p$ periphery banks of investment size $i_p \ll 1$, and each core bank is linked to k_p periphery banks with offsetting risks of standard deviation $\sigma_p = \sqrt{2}\sigma \cdot i_p$. The resulting OTC network structure is depicted in Figure 5, and the risk of each core bank is scaled to the baseline model: $\sqrt{2}\sigma = \sqrt{2\sigma_c^2 + k_p\sigma_p^2}$. There are two kinds of separate exogenous crisis events, thus three states of nature. There is a “small” crisis with probability p_p : one periphery bank of each core bank, selected at random, fails at $t = 1$. With probability $p_c < p_p$ there is a “big” crisis in which not only do n periphery banks fail but also one core bank, selected at random.

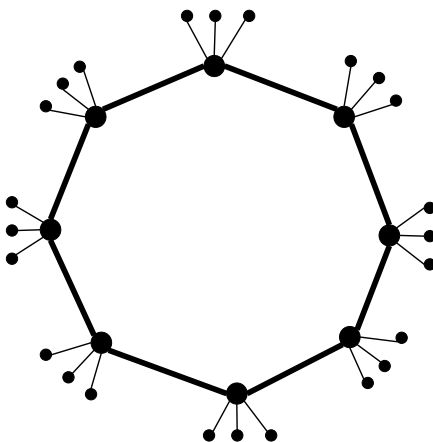


Figure 5: The full financial system

A stylized graph of the full financial system analyzed in Subsection 5.3 with $n = 6$ and $k_p = 3$. While large banks in the core are linked in a circle through OTC contracts, each small periphery bank contracts with one large core bank.

I propose an equilibrium, for some crisis probabilities, in which core banks hold enough capital to survive the collapse of a periphery bank, but do not insure against the failure of a core bank. Thus the system is resilient to small crises but collapses in a big crisis. Note that since one periphery bank fails at each core bank in a small crisis, there is no scope for mutual insurance between core banks against the failure of periphery banks. The core banks keep capital to survive the failure of one of

their periphery banks if and only if:

$$\begin{aligned} & (1 - p_c) \cdot R + p_c \cdot L - 1 - \underline{K}(0) \cdot c - [\underline{K}(\sigma_p) - \underline{K}(0)] \cdot c \\ & > (1 - p_c - p_p) \cdot R + (p_c + p_p) \cdot L - 1 - \underline{K}(0) \cdot c \end{aligned} \quad (33)$$

if σ_p is not too large and p_p not too small, then the condition below is satisfied:

$$p_p > p_p^i = c \cdot \frac{\underline{K}(\sigma_p) - \underline{K}(0)}{R - L} \quad (34)$$

I assume in the following that $p_p > p_p^i$, thus core banks prepare for the default of periphery banks.

Following the derivation of Equation 6, and recognizing that it takes capital of $\underline{K}(\sqrt{\sigma_c^2 + \sigma_p^2}) - \underline{K}(\sigma_p)$ to stabilize each core counterparty of a failing core bank, the counterparty insurance scheme against the failure of other core banks is sustainable only if:

$$p_c > p_{c,out}^i = c \cdot \frac{\underline{K}(\sqrt{\sigma_c^2 + \sigma_p^2}) - \underline{K}(\sigma_p)}{R - L} \quad (35)$$

On the other hand, the social planner would insure if and only if:

$$\begin{aligned} & (1 + k_p \cdot i_p) \cdot [(n - p_c) \cdot R + p_c \cdot L - n] - n \cdot c \cdot \underline{K}(\sigma_p) - c \cdot 2 \cdot [\underline{K}(\sqrt{\sigma_c^2 + \sigma_p^2}) - \underline{K}(\sigma_p)] \\ & > (1 + k_p \cdot i_p) \cdot [(n - n \cdot p_c) \cdot R + n \cdot p_c \cdot L - n] - n \cdot c \cdot \underline{K}(\sigma_p) \end{aligned} \quad (36)$$

where I exclude the cost of capital to the periphery banks, since it would cancel out. Thus the counterparty insurance between core banks is socially optimal if:

$$p_c > p_c^s = \frac{2}{n - 1} \cdot c \cdot \frac{\underline{K}(\sqrt{\sigma_c^2 + \sigma_p^2}) - \underline{K}(\sigma_p)}{(R - L) \cdot (1 + k_p \cdot i_p)} \quad (37)$$

For the full financial system the inefficiency region is even wider than that for the core only, at least in a relatively sense: the ratio of the crisis probability at which counterparty insurance is socially optimal to that for which it is privately optimal is smaller. The ratio is not simply $\frac{2}{n-1}$, as in case of no periphery in Equations 6 and 8, but $\frac{2}{n-1} \cdot \frac{1}{1+k_p \cdot i_p}$. The reason is that core banks do not simply exert an externality on other core banks, but also on all periphery banks in the whole system.

This example also shows that core banks do hold some slack equity in normal times, they do not exactly operate at their capital constraint. However, this small capital cushion is set aside for the failure of small counterparties. This does not contradict the fact that core banks do not insure against the failure of other core banks.

6 Discussion of the results and the crisis of 2008

There are three questions to answer in order to assess whether the mechanism shown in this paper is realistic in a modern financial system, in particular whether it could have contributed to the credit market freeze during the financial crises of 2008. First, I consider whether each bank has only a few large counterparties. Second, I analyze whether the failure of a financial institution could lead to a run on its major counterparties. Third, I give anecdotal evidence that financial institutions do not necessarily insure against the loss of counterparty.

The results of the model show that the crucial question when deciding whether the financial system is likely to be inefficient is the density of the hedging network. Thus the regulator should collect information on the size of OTC hedging exposure of banks to assess the stability of the financial system. Unfortunately, the data on OTC 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 OTC hedging contracts. The payments to counterparties are a good proxy of the size of OTC exposures. Table 1 shows that these payments were heavily concentrated towards a few banks. The concentration of payments is strong: the top three counterparties account for about half of the payments, hinting at a sparse network.

There is also indirect evidence that OTC exposures could lead to contagion due to the loss of hedging contracts. First, Moody's has claimed that replacement costs incurred by Lehman counterparties was a strong indirect contagion channel around the Lehman bankruptcy.¹³ Second, according to Table 1, some exposures to AIG were large enough to potentially prompt a run of wholesale creditors on its counterparties. For example, Goldman received payments of 11 billion by the end of the year 2008 from AIG for OTC hedges after AIG 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 would have incurred substantial equity losses. While Goldman claimed after the crisis that it had ample insurance against the default of AIG,¹⁴ this claim could not be verified by investors and it is disputed whether it would have covered the losses.¹⁵ My model shows that it was reasonable for investors to believe that Goldman might not have had the proper incentives to hedge.

The events around the default of Lehman Brothers yields some evidence that counterparties did not have enough insurance ex ante against the default of the banks they did business with. The spike in CDS prices of major OTC counterparties can be at least partially attributed to counterparties trying to hedge their exposure and thus driving up spreads.¹⁶

¹³ "Credit Default Swaps and Counterparty Risk", European Central Bank, August 2009

¹⁴ "Goldman Confirms \$6 Billion AIG Bets", The Wall Street Journal, March 21, 2009

¹⁵ "Goldman's Price of Protection", The Wall Street Journal, March 18, 2009

¹⁶ "Lehman's Demise Triggered Global Cash Crunch", The Wall Street Journal, September 29, 2008

7 Concluding remarks

This paper develops a model of an entangled financial system where banks use OTC contracts to hedge their risks. In such a system, the failure of a single bank can lead to the complete collapse of the financial system as creditors run on all banks simultaneously. I show that even though banks use OTC contracts to hedge risks and thereby expose themselves to counterparty risk, they are unwilling to insure against counterparty default. The reason is a market failure: the externalities of bankruptcy inflicted on others through derivative contracts is not internalized by the banks. The paper also takes a step in clarifying the effect of entanglement: while both no entanglement and full entanglement lead to safe systems, it is partially entangled systems that are most prone to crisis and inefficiency.

Based on limited data and anecdotal evidence, I draw the conclusion that the conditions for such a contagious equilibrium are realistic. Thus the results of the paper can be used to guide the regulator and further empirical research on what kind of data to gather and to test whether the underlying assumption are indeed met in the financial system. The framework of the model can also be used to analyze other questions in “entangled” financial systems. In Zawadowski (2009a) I extend the analysis to a developing financial system and show that the conditions for inefficiency are most likely to be met during the initial boom in OTC contracting. Furthermore, the framework can be used to analyze the systemic effects of margin accounts. Beyond the financial sector, the insight regarding the low ex ante choice of counterparty insurance can be extended to modeling industrial networks and trade.

Appendix: Proofs

Lemma 1

Proof. 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 $D_i = 1 + G_i - K_i$ this becomes: $(R_{f,1} - R_{l,1}R_{l,0}) \cdot G_i - R_{l,1}R_{l,0} \cdot (1 - K_i)$

Given that the probability of default of any institution is non-negative, the nominal interest rate on the loan must be higher than the risk-free rate, implying: $R_{l,1}R_{l,0} \geq R_{f,1}$. Since survival at $t = 1$ also depends on the expected payoff at $t = 2$, it also is increasing in G_i . Thus choosing $G_i = 0$ is preferable in terms of payoff. Since the incentive constraint and the participation constraint also depend on the final payoff in the same way as the objective function, choosing $G_i = 0$ is optimal. \square

Lemma 2

Proof. The incentive constraint of a bank surviving at $t = 1$ can be rewritten as following:

$$R - R_{f,1}R_{l,0} \cdot D + (1 - \pi_{2|1}(R_{l,0}, \sigma_1)) \cdot X \geq B_1 \quad (38)$$

The reason is the following. The pledgable payoff of the project is R independent of the debt level. The expected payoff of the investors to make them break 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 $t = 1$: $[1 - \pi_{2|1}(R_{l,0}, R_{l,1}, D)] \cdot X$. 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.

The bank chooses $K = 1 - D$, s.t. the incentive constraint binds: this yields implicit equation 20 for the equity. Restriction 18 implies that $\underline{K}(R_{l,0}, \sigma_1) > 0$ 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 $R_{l,0}\underline{D} = R_{l,0}(1 - \underline{K})$ at $t = 1$:

$$R_{f,1}R_{l,0}\underline{D} = \underline{R}_{l,1}R_{l,0}\underline{D} - \int_{-\infty}^{\underline{R}_{l,1}R_{l,0}\underline{D}} (\underline{R}_{l,1}R_{l,0}\underline{D} - x) \frac{1}{\sigma_1} \phi\left(\frac{x - R}{\sigma_1}\right) dx \quad (39)$$

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 implicit expression for $\underline{R}_{l,1}(R_{l,0}, \sigma_1)$:

$$(\underline{R}_{l,1} - R_{f,1})R_{l,0}\underline{D} = \sigma_1 \cdot \phi\left(\frac{\underline{R}_{l,1}R_{l,0}\underline{D} - R}{\sigma_1}\right) - (R - \underline{R}_{l,1}R_{l,0}\underline{D}) \cdot \Phi\left(\frac{\underline{R}_{l,1}R_{l,0}\underline{D} - R}{\sigma_1}\right) \quad (40)$$

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

$$\frac{\partial \underline{K}}{\partial \sigma_1} = \frac{X}{R_{l,0}R_{f,1}} \cdot \frac{\Phi(x) \cdot (R - R_{l,0}\underline{R}_{l,1}(1 - \underline{K})) + \phi(x)\sigma}{\sigma^2 \frac{\Phi(x)}{\phi(x)} + \sigma X} > 0 \quad (41)$$

where $(R - R_{l,0}\underline{R}_{l,1}(1 - \underline{K})) > 0$ 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. \square

Lemma 3

Proof. A bank goes 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 $\underline{D}^h(R_{l,0}) = 1 - \underline{K}^h(R_{l,0})$, the condition for bankruptcy is $R - B_1 + X > R_{f,1}L$, which in turn is satisfied by restriction 19. 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 restriction 16, i.e. that $R - d < L$, and the investors optimally liquidate the real investment. \square

Lemma 4

Proof. If neither of the counterparties default, exerting effort is optimal, since the incentive constraint is satisfied with equity $\underline{K}^h(R_{l,0})$. If one or two of the counterparties defaults, the risk of payoff (measured as of $t = 1$) σ_1 increases. By Lemma 2 $\frac{\partial \underline{K}}{\partial \sigma_1} > 0$, 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. Also, the 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 $t = 1$ either: the incentive constraint is only effected by the change in the expected value of non-pledgable payoff. \square

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, $K = 0$ 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 restriction 18 is lower than $B_0 + B_1$. Thus the bank has no incentive to exert effort with long-term debt, short-term debt is used in equilibrium.

We also have to check whether short-term debt is feasible. Choosing short-term debt the bank can still choose to hold $K = 0$, 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:

$$E[P] = (1-p) \cdot [R + X - R_{f,1}R_{l,0}\underline{D}^h(R_{l,0})] - R_e\underline{K}^h(R_{l,0}) = (1-p)B_1 - R_e \cdot \left(\frac{B_1 - R - X}{R_{f,1}R_{l,0}^h} + 1 \right) \quad (42)$$

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:

$$(1-p) \cdot B_1 - R_e \left((1-p) \cdot \frac{B_1 - R - X}{R_{f,1}} + 1 \right) > B_0 \quad (43)$$

which is satisfied by Assumption 1. \square

Lemma 6

Proof. The first potential gain from not hedging ϵ risks is that equityholders can shift risk to debtholders. 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. The second possible gain from not hedging ϵ risks that the bank avoids contagion in crisis. When the bank's counterparty defaults on its OTC 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. All in all, there cannot be any gain from not hedging ϵ risks, furthermore not hedging ϵ risk increases the probability of bankruptcy at $t = 2$, thus decreasing the expected value of non-pledgable payoff, implying that banks are not simply indifferent between hedging and not, they strictly prefer to hedge. \square

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 and only if:

$$\begin{aligned} & (1-p) \cdot [R - R_{f,1}R_{l,0}^h\underline{D}^h(R_{l,0}^h) + X] - R_e\underline{K}^h(R_{l,0}^h) \\ & \geq (1-p) \cdot [R - R_{f,1}R_{l,0}^h\underline{D}^a(R_{l,0}^h) + X] + \frac{n-1}{n}p \cdot [R - R_{f,1}R_{l,0}^h\underline{D}^a(R_{l,0}^h) + (1 - \tilde{\pi}_{2|1})X] - R_e\underline{K}^a(R_{l,0}^h) \end{aligned} \quad (44)$$

where

$$\tilde{\pi}_{2|1} = \Phi \left(\frac{\tilde{R}_{l,1}R_{l,0}^h\underline{D}^a(R_{l,0}^h) - R}{\sqrt{2}\sigma} \right) \quad (45)$$

The left hand side of the inequality is the expected payoff the contagious system, and the right hand side is that when the bank holds enough equity $\underline{K}^a(R_{l,0}^h)$ to survive a crisis. Note that since investors cannot make the ex ante interest rate conditional on equity choice, they charge the same initial interest rate to the bank that chooses to deviate and hold

higher equity. A bank with equity $\underline{K}^a(R_{l,0}^h)$ 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 case of a systemic crisis. In a systemic crisis the deviating bank is risky, even though it can roll over its debt, so the interest rate on its loan will jump to $\tilde{R}_{l,1} > R_{f,1}$, which is set s.t. the investors break even. Rearranging and using that the incentive constraint at $t = 1$ is binding yields:

$$\begin{aligned} R_e[\underline{K}^a(R_{l,0}^h) - \underline{K}^h(R_{l,0}^h)] + (1-p) \cdot \underbrace{R_{f,1}[R_{l,0}^h \underline{D}^a(R_{l,0}^h) - R_{l,0}^h \underline{D}^h(R_{l,0}^h)]}_{-\tilde{\pi}_{2|1}X} \\ \geq \frac{n-1}{n}p \cdot \underbrace{[R - R_{f,1}R_{l,0} \underline{D}^a(R_{l,0}^h) + (1 - \tilde{\pi}_{2|1})X]}_{B_1} \end{aligned} \quad (46)$$

rearranging we arrive at the implicit Equation 22 for p^a .

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. I take the partial derivatives of equations 45 and 40. This gives a system of two equations with two unknowns: $\frac{\partial \pi_{2|1}}{\partial K}$ and $\frac{\partial \tilde{R}_{l,1}}{\partial K}$. We solve for $\frac{\partial \pi_{2|1}}{\partial K}$: the first equation is the marginal increase in survival probability at $t = 2$, in case more equity is held ex ante. The second is the expected payoff given equity K :

$$\begin{aligned} \frac{\partial \pi_{2|1}}{\partial K} &= -\frac{1}{\sigma_1} \cdot \frac{\phi\left(\frac{R - \tilde{R}_{l,1} R_{l,0}^h (1-K)}{\sigma_1}\right)}{1 - \Phi\left(\frac{R - \tilde{R}_{l,1} R_{l,0}^h (1-K)}{\sigma_1}\right)} \cdot R_{f,1} R_{l,0}^h \\ E[P] &= (1-p) \cdot [R - R_{f,1} R_{l,0}^h \cdot (1-K) + X] \\ &+ \frac{n-1}{n}p \cdot [R - R_{f,1} R_{l,0}^h \cdot (1-K) + (1 - \pi_{2|1}(K)) \cdot X] - R_e K \end{aligned} \quad (47)$$

where the second term in Equation 48 is the payoff in case the whole financial system collapses except for the one holding surplus equity. The change in $E[P]$ if the bank increases K over the minimum level needed for rollover:

$$\left. \frac{\partial E[P]}{\partial K} \right|_{K=K^a(R_{l,0}^h)} = \left(1 - \frac{p}{n}\right) \cdot R_{f,1} R_{l,0}^h - \frac{n-1}{n}pX \cdot \frac{\partial \pi_{2|1}}{\partial K} - R_e \quad (49)$$

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:

$$\frac{n-1}{n}pX \cdot \left(-\frac{\partial \pi_{2|1}}{\partial K}\right) \leq \frac{p}{n}R_{f,1}R_{l,0}^h + (R_e - R_{f,1}R_{l,0}^h) \quad (50)$$

Substituting $\frac{\partial \pi_{2|1}}{\partial K}$, it is sufficient to show that:

$$pX \cdot \frac{1}{\sigma_1} \cdot \frac{\phi\left(\frac{R - \tilde{R}_{l,1} R_{l,0}^h (1-K)}{\sigma_1}\right)}{1 - \Phi\left(\frac{R - \tilde{R}_{l,1} R_{l,0}^h (1-K)}{\sigma_1}\right)} \leq \frac{R_e}{R_{f,1}} \left(1 - \frac{p}{n}L\right) - 1 \quad (51)$$

Given that $\frac{\phi(x)}{1-\Phi(x)} \leq \sqrt{\frac{2}{\pi}}$ for any x , and $\sigma_1 = \sqrt{2}\sigma$, this holds by Assumption 2. 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:

$$1 \cdot \underline{D}^h(R_{l,0}) = (1-p) \cdot R_{l,0} \underline{D}^h(R_{l,0}) + p \cdot L \quad (52)$$

Substituting $\underline{D}^h(R_{i,0}) = \frac{R+X-B_1}{R_{f,1}R_{i,0}^h}$ 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. The complete collapse of the network in crisis gives a lower bound on profits. Given that the incentive constraint is binding, the expected profit is:

$$E[P] \geq (1-p)[R+X - R_{f,1}R_{i,0}^h\underline{D}^h(R_{i,0}^h)] - R_e\underline{K}^h(R_{i,0}^h) = (1-p)B_1 - R_e \cdot \left(\frac{B_1 - R - X}{R_{f,1}R_{i,0}^h} + 1 \right) \quad (53)$$

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

$$(1-p)B_1 - R_e \left(1 - (1-p) \frac{R+X-B_1}{R_{f,1}} \right) > 0 \quad (54)$$

which is satisfied by Assumption 1, given that $B_0 \geq 0$. Thus entrepreneurs choose to participate. \square

Lemma 7

Proof. 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 $K = K^h(R_{i,0})$ it can achieve payoff of $B_0 + B_1 - R_e \cdot K^h(R_{i,0})$. On the other hand if it chooses high effort, the upper bound on its profits is: $R - R_{f,1} + X - R_e \cdot K^h(R_{i,0})$ where we assumed it could borrow all fund for the investment at the riskless rate, and that the investment succeed with probability one. Clearly, by restriction 18, it is a profitable deviation to shirk since $B_0 > 0$. Thus insuring against $s_i = d$ is not feasible. \square

Proposition 2

Proof. First, I simply assume that the counterparties of the failing bank get only enough equity to roll over debt: I show that this is indeed the case at the end of this proof. Thus the insurance pool needs reserves adding up to $2[\underline{K}^u(R_{i,0}) - \underline{K}^h(R_{i,0})]$ to stabilize the two counterparties of the failing bank. 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:

$$\begin{aligned} & (1 - 3\frac{p}{n}) [R - R_{f,1}\hat{R}_{i,0}\underline{D}^h(\hat{R}_{i,0}) + X] - R_e\underline{K}^h(\hat{R}_{i,0}) \geq \\ & \geq (1 - \frac{p}{n}) \left[R - R_{f,1}\hat{R}_{i,0}\underline{D}^h(\hat{R}_{i,0}) + X \right] - R_e\underline{K}^h(\hat{R}_{i,0}) + R_{f,1} \cdot \underbrace{\frac{2p}{n} (\underline{K}^u(\hat{R}_{i,0}) - \underline{K}^h(\hat{R}_{i,0}))}_{\text{expected insurance receivable}} - 2\frac{p}{n}\hat{\pi}_{2|1}X - k \end{aligned} \quad (55)$$

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

$$\hat{\pi}_{2|1} = \Phi \left(\frac{\hat{R}_{i,1}\hat{R}_{i,0}\underline{D}^u - R}{\sigma} \right) \quad (56)$$

and $\hat{R}_{i,0}$ is the endogenous short-term rate if all banks choose to insure and $\hat{R}_{i,1}$ is the interest rate charged to the two banks that get capital injection in the crisis state at $t = 1$. If the bank is uninsured but all the others are insured, it goes bankrupt if either it or its two counterparties 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 $-2\frac{p}{n}\hat{\pi}_{2|1}X$ has to be added, since even if a bank's counterparty failed and it has $\underline{K}^u(\hat{R}_{l,0})$, i.e. can survive at $t = 1$, it may still fail at $t = 2$ and lose the non-pledgable payoff. Note that 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. Also, having extra equity funds in the insurance vehicle does not allow the banks to borrow more, since in crisis these reserves are all transferred to the banks neighboring the one in trouble and all other banks have to be able to roll over their debt given 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:

$$k = \frac{1}{n} \left[2 \left(\underline{K}^u(\hat{R}_{l,0}) - \underline{K}^h(\hat{R}_{l,0}) \right) (R_e - R_{f,1}) + R_{f,1} \cdot 2p \left(\underline{K}^u(\hat{R}_{l,0}) - \underline{K}^h(\hat{R}_{l,0}) \right) \right] \quad (57)$$

where $2p \left(\underline{K}^u(\hat{R}_{l,0}) - \underline{K}^h(\hat{R}_{l,0}) \right)$ is the expected payout of the insurance fund. We arrive at an implicit equation for p^i , which yields Equation 25:

$$\hat{p}^i : \quad p = \frac{\left(\underline{K}^u(\hat{R}_{l,0}) - \underline{K}^h(\hat{R}_{l,0}) \right) (R_e - R_{f,1})}{R - R_{f,1} \hat{R}_{l,0} \underline{D}^h(\hat{R}_{l,0}) + X - \hat{\pi}_{2|1} X} = \frac{\hat{\pi}_{2|1} X (R_e - R_{f,1})}{R_{f,1} \hat{R}_{l,0} (B_1 - \hat{\pi}_{2|1} X)} \quad (58)$$

Now I calculate social welfare. 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 welfare per entrepreneur is:

$$W = (1 - \pi_1) \cdot [R + X - R_{f,1} R_{l,0} D] - \pi_{2|1} X - R_e K \quad (59)$$

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 $\pi_{2|1}$, it loses the non-pledgable payoff. Using 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 $D = 1 - K$ on the left hand side and rearranging this yields:

$$(1 - \pi_1) R_{f,1} R_{l,0} D = R_{f,1} - R_{f,1} K - \pi_1 R_{f,1} L \quad (60)$$

which we now substitute into the welfare W :

$$W = (1 - \pi_1) \cdot (R + X) - \pi_{2|1} X + \pi_1 R_{f,1} L - R_{f,1} - (R_e - R_{f,1}) K \quad (61)$$

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 counterparty of the failed bank, times the conditional probability of going bankrupt with equity $\underline{K}^u(\hat{R}_{l,0})$, thus $\pi_{2|1} = 2\frac{p}{n}\hat{\pi}_{2|1}$. The social benefits outweigh the social costs of the counterparty insurance for the system as a whole if and only if:

$$n(1 - p)[R + X] + npR_{f,1}L \leq (n - p)[R + X] + pR_{f,1}L - 2p\hat{\pi}_{2|1}X - 2 \left(\underline{K}^u(\hat{R}_{l,0}) - \underline{K}^h(\hat{R}_{l,0}) \right) \cdot (R_e - R_{f,1}) \quad (62)$$

where we simplified by $n\underline{K}^h(\hat{R}_{l,0}) \cdot (R_e - R_{f,1})$. The left hand side is the total welfare without insurance, while the right hand with insurance. Thus counterparty insurance is socially optimal if $p > p^s$, where the implicit equation for p^s is:

$$p^s : \quad p = \frac{2}{n-1} \frac{\left(\underline{K}^u(\hat{R}_{l,0}) - \underline{K}^h(\hat{R}_{l,0}) \right) (R_e - R_{f,1})}{R + X - R_{f,1}L - \frac{2}{n-1}\hat{\pi}_{2|1}X} \quad (63)$$

where substituting $\underline{K}^u(\hat{R}_{l,0}) - \underline{K}^h(\hat{R}_{l,0}) = \frac{\hat{\pi}_{2|1}X}{R_{f,1}\hat{R}_{l,0}}$ yields Equation 24.

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 nominal 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 assumed that banks getting capital injection from the insurance fund receive exactly enough to roll over their debt at $t = 1$. Here I show that banks do not choose to transfer more equity than that to the two banks. The expected payoff for an individual bank under the insurance scheme is:

$$E[P] = \left(1 - \frac{p}{n}\right) \left[R - R_{f,1}\hat{R}_{l,0}\underline{D}^h(\hat{R}_{l,0}) + X \right] - R_e\underline{K}^h(\hat{R}_{l,0}) + (R_e - R_{f,1}) \cdot \frac{2}{n} \left(\underline{K}^u(\hat{R}_{l,0}) - \underline{K}^h(\hat{R}_{l,0}) \right) - 2\frac{p}{n}\hat{\pi}_{2|1}X \quad (64)$$

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

$$\frac{\partial E[P]}{\partial \underline{K}^u} = -\frac{2}{n} \cdot (R_e - R_{f,1}) - \frac{2p}{n}X \frac{\partial \hat{\pi}_{2|1}}{\partial \underline{K}^u} \quad (65)$$

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, :

$$pX \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} \leq \left(\frac{R_e}{R_{f,1}} - 1 \right) \cdot \left(1 - \frac{p}{n}L \right) \quad (66)$$

which is fulfilled by Assumption 2. □

References

- Acharya, Viral and Alberto Bisin**, “Centralized versus Over-The-Counter Markets,” *New York University, Working Paper*, 2009.
- , **Denis Gromb, and Tanju Yorulmazer**, “Imperfect Competition in the Interbank Market for Liquidity as a Rationale for Central Banking,” *London Business School, Working Paper*, 2008.
- , **Irind Gujral, and Hyun Song Shin**, “Dividends and Bank Capital in the Financial Crisis of 2007-2008,” *New York University, Working Paper*, 2009.
- Adrian, Tobias and Hyun Song Shin**, “Liquidity and financial contagion,” *Banque de France Financial Stability Review*, 2008, 81 (11), 1–7.
- and —, “Liquidity and leverage,” *Journal of Financial Intermediation*, forthcoming.

- Akerlof, George A.**, “The Market for ”Lemons”: Quality Uncertainty and the Market Mechanism,” *Quarterly Journal of Economics*, 1970, 84 (3), 488–500.
- Allen, Franklin and Douglas Gale**, “Financial Contagion,” *Journal of Political Economy*, 2000, 108 (1), 1–33.
- Babus, Ana**, “The Formation of Financial Networks,” *Cambridge University, Working Paper*, 2009.
- Bech, Morten L. and Enghin Atalay**, “The Topology of the Federal Funds Market,” *Federal Reserve Bank of New York, Staff Report No. 354*, 2008.
- Bergman, William J, Robert R. Bliss, Christian A. Johnson, and George G. Kaufman**, “Netting, Financial Contracts, and Banks: The Economic Implications,” in George G. Kaufman, ed., *Market Discipline in Banking: Theory and Evidence*, Elsevier, 2003, pp. 303–334.
- 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.
- Bliss, Robert R. and George G. Kaufman**, “Derivatives and Systemic Risk: Netting, Collateral, and Closeout,” *Journal of Financial Stability*, 2006, 2, 55–70.
- 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.
- and **Martin Oehmke**, “The Maturity Rat Race,” *Princeton University, Working Paper*, 2009.
- Caballero, Ricardo J. and Alp Simsek**, “Complexity and Financial Panics,” *NBER working paper, No. 14997*, 2009.
- Calomiris, Charles W.**, *U.S. Bank Deregulation in Historical Perspective*, Cambridge University Press, 2000.
- and **Charles M. Kahn**, “The Role of Demandable Debt in Structuring Optimal Banking Arrangements,” *American Economic Review*, 1991, 81 (3), 497–513.
- Cifuentes, Rodrigo, Gianluigi Ferrucci, and Hyun Song Shin**, “Liquidity Risk and Contagion,” *Journal of the European Economic Association*, 2005, 3 (2-3), 556–566.
- Diamond, Douglas W. and Philip H. Dybvig**, “Bank Runs, Deposit Insurance and Liquidity,” *Journal of Political Economy*, 1983, 91 (3), 401–419.
- and **Raghuram G. Rajan**, “Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking,” *Journal of Political Economy*, 2001, 109 (2), 287–327.
- Duffie, Darrell and Haoxiang Zhu**, “Does a Central Clearing Counterparty Reduce Counterparty Risk?,” *Stanford University, Working Paper*, 2009.

- Furfine, Craig**, “Interbank Exposures: Quantifying the Risk of Contagion,” *American Economic Review Papers and Proceedings*, 2003, 35 (1), 111–128.
- Gai, Prassana and Sujit Kapadia**, “Contagion in Financial Networks,” *Bank of England, Working Paper*, 2007.
- Gorton, Gary and George Pennacchi**, “Financial Intermediaries and Liquidity Creation,” *Journal of Finance*, 1990, 45 (1), 49–71.
- He, Zhiguo and Wei Xiong**, “Dynamic Debt Runs,” *Princeton University, Working Paper*, 2009.
- Holmström, Bengt and Jean Tirole**, “Financial Intermediation, Loanable Funds, and the Real Sector,” *Quarterly Journal of Economics*, 1997, 112, 663–692.
- Innes, R.**, “Limited liability and incentive contracting with ex ante action choices,” *Journal of Economic Theory*, 1990, 75, 45–67.
- Jackson, Matthew O. and Asher Wolinsky**, “A Strategic Model of Social and Economic Networks,” *Journal of Economic Theory*, 1996, 71, 44–74.
- 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.
- Lagunoff, Roger and Stacey L. Schreft**, “A Model of Financial Fragility,” *Journal of Economic Theory*, 2001, 99, 220–264.
- Leitner, Yaron**, “Financial Networks: Contagion, Commitment, and Private Sector Bailouts,” *Journal of Finance*, 2005, 60 (6), 2925–2953.
- Lorenzoni, Guido**, “Inefficient Credit Booms,” *Review of Economic Studies*, 2008, 75 (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.
- Rampini, Adriano A. and S. Viswanathan**, “Collateral, Risk Management, and the Distribution of Debt Capacity,” *Duke University, Working Paper*, 2009.
- Rotemberg, Julio J.**, “Liquidity Needs in Economies with Interconnected Financial Obligations,” *NBER working paper, No. 14222*, 2008.
- Shleifer, Andrei and Robert W. Vishny**, “The Limits of Arbitrage,” *Journal of Finance*, 1997, 52 (1), 35–55.
- Soramäki, Kimmo, Morten L. Bech, Jeffrey Arnold, Robert J. Glass, and Walter E. Beyeler**, “The Topology of Interbank Payment Flows,” *Physica A*, 2007, 379, 317–333.
- Stulz, René M.**, “Credit Default Swaps and the Credit Crisis,” *NBER Working Paper 15384*, 2009.

Upper, Christian and Andreas Worm, “Estimating bilateral exposures in the German interbank market: Is there a danger of contagion?,” *European Economic Review*, 2004, 48, 827–849.

Zawadowski, Adam, “The Development of OTC Markets and Crises,” *Princeton University, Working Paper*, 2009.

—, “Liquidity Hoarding in Credit Networks,” *Princeton University, Working Paper*, 2009.

Tables

counterparty	Payments (billions)			cumulative total
	CDS coll.	Maiden III	total	
Société Général	4.1	6.9	11.0	
Goldman Sachs	2.5	5.6	8.1	
Deutsche Bank	2.6	2.8	5.4	49.5%
Merrill Lynch	1.8	3.1	4.9	59.4%
UBS	0.8	2.5	3.3	66.1%
Calyon	1.1	1.2	2.3	70.7%
Barclays	0.9	0.6	1.5	73.7%
Bank of America	0.2	0.5	0.7	75.2%
total	22.4	27.1	49.5	100.0%

Table 1: Payments of AIG to top counterparties: September 16 to December 31, 2008.

The first column are the margin payments for credit default swaps, the second payments through the limited liability company Maiden III, in billions of US dollars. The last column is the cumulative payments to top counterparties and their proportion as of all payments. Payments due to securities shorting are not included. Source: AIG, March 15, 2009, from Attachments A, B, and D. Available online (last accessed September 21, 2009): www.aig.com/aigweb/internet/en/files/CounterpartyAttachments031809_tcm385-155645.pdf